
Volume 95
Number 2
2008

MISSOURI BOTANICAL
JUN 24 2008
GARDEN LIBRARY

Annals
of the
Missouri
Botanical
Garden



ASSUMPTION 0 ANALYSIS:
COMPARATIVE PHYLOGENETIC
STUDIES IN THE AGE OF
COMPLEXITY^{1, 2, 3}

Daniel R. Brooks⁴ and Marco G. P. van Veller⁵

ABSTRACT

Darwin's panoramic view of biology encompassed two metaphors: the phylogenetic tree, pointing to relatively linear (and divergent) complexity, and the tangled bank, pointing to reticulated (and convergent) complexity. The emergence of phylogenetic systematics half a century ago made it possible to investigate linear complexity in biology. Assumption 0, first proposed in 1986, is not needed for cases of simple evolutionary patterns, but must be invoked when there are complex evolutionary patterns whose hallmark is reticulated relationships. A corollary of Assumption 0, the duplication convention, was proposed in 1990, permitting standard phylogenetic systematic ontology to be used in discovering reticulated evolutionary histories. In 2004, a new algorithm, phylogenetic analysis for comparing trees (PACT), was developed specifically for use in analyses invoking Assumption 0. PACT can help discern complex evolutionary explanations for historical biogeographical, coevolutionary, phylogenetic, and tokogenetic processes.

Key words: Assumption 0, biogeography, cladogeny, coevolution, evolutionary complexity, horizontal transfer, hybridization, PACT, phylogenetic analysis for comparing trees, phylogenetic systematics, phylogeography, reticulation, taxon pulse, tokogeny.

¹This and the three articles that follow it are the proceedings of the 52nd Annual Systematics Symposium of the Missouri Botanical Garden, "Reconstructing Complex Evolutionary Histories: Gene-Species Trees, Historical Biogeography, and Coevolution." The symposium was held 7–9 October 2005, at the Missouri Botanical Garden in St. Louis, Missouri, U.S.A.

²This was the 50th Missouri Botanical Garden Annual Systematics Symposium to be supported by a grant from the National Science Foundation (grant DEB-0515933). The meeting was organized and moderated by Richard Mayden of Saint Louis University; P. Mick Richardson was responsible for the smooth running of the symposium, with the welcome and able assistance of Alina Freire-Fierro, Mary McNamara, Sandy Lopez, Donna Rodgers, William Guy, and Zubin Chandran; Victoria C. Hollowell, Beth Parada, and Allison Brock (Missouri Botanical Garden Press) were responsible for the publication of the symposium proceedings.

³This contribution is based on a presentation made by DRB at the 2005 Annual Systematics Symposium of the Missouri Botanical Garden. We thank Rick Mayden for the invitation to participate in that symposium, as well as Peter Raven, Mick Richardson, and the other organizers of the symposium. We also express special appreciation to Marco Roos, E. O. Wiley, and Rino Zandee, who first articulated Assumption 0. DRB acknowledges funding from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

⁴Department of Zoology, University of Toronto, 25 Harbord Street, Toronto, Ontario M5S 3G5, Canada. dbrooks@zoo.utoronto.ca.

⁵Biosystematics Group, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands. marco.vanveller@wur.nl.

doi: 10.3417/2006017

The universe is structured by laws, and science is the search for methods of data analysis and theories providing powerful general explanations, all in terms of those general laws. This ontology of simplicity has guided the development of western science for nearly 2500 years, embodied in the principle of parsimony (Latin “*parcere*,” to spare). Aristotle (350 B.C.E.) articulated the ontological view of the principle of parsimony, the postulate that “nature operates in the shortest way possible” and “the more limited, if adequate, is always preferable.” The principle of parsimony is also linked with the English philosopher and Franciscan monk William of Ockham (ca. 1285–1349), who advocated the use of what is known as “Ockham’s razor”: “*pluralitas non est ponenda sine neccesitate*” (“plurality should not be posited without necessity”) and “*non sunt multiplicanda entia praeter necessitatem*” (“entities should not be multiplied unnecessarily”). In this sense, the principle of simplicity obliges us to favor theories or hypotheses that make the fewest unwarranted, or ad hoc, assumptions about the data from which they are derived. This epistemological use of the principle does not necessarily imply that nature itself is parsimonious. Indeed, despite the best efforts of philosophers for more than 700 years, no link between parsimony and truth has ever been established. Nonetheless, most scientists conduct their research as if they believe that nature is parsimonious in some sense and, as a consequence, they place much more credence in simple than in complex theories.

The cosmologist Stephen Hawking has dubbed the 21st century the century of complexity, leading a parade of physicists who began to think about ontological complexity in the latter decade of the 20th century. Biology should be well poised to take advantage of this sea change in our understanding of the nature of science and the universe, having discovered an ontology of complexity called Darwinism almost 150 years ago. Few fully realize that Darwinism is not a simple theory:

“... there are two factors: namely, the nature of the organism and the nature of the conditions. The former seems to be much more the important; for nearly similar variations sometimes arise under, as far as we can judge, dissimilar conditions; and, on the other hand, dissimilar variations arise under conditions which appear to be nearly uniform.”

—Darwin (1872: 32)

Most evolutionary biologists consider this passage no more than a general repudiation of Lamarckism. We believe, however, that it is one of the first modern articulations of a complex theory: Darwin proposed that evolution is an emergent property of asymmetrical interactions between two different causal agents, each

with its own properties, producing outcomes that are not readily predictable from knowledge of the properties of either agent and, thus, historically contingent.

Darwin thought that organisms were historically and developmentally cohesive wholes, and it was in the “nature of the organism” to produce offspring that were all highly similar (but not identical) to each other and to their parents and other ancestors. He also postulated that reproduction produced variation without regard for environmental conditions, and it was in the “nature of the organism” to produce these offspring in numbers far exceeding the resources available for their support. When this inherent overproduction produced variety in critical characters, natural selection would preserve the versions that were functionally superior in that particular environmental context (adaptations). Whenever an environment changed, those organisms that already had the adaptations necessary to survive would do so, whereas those lacking appropriate adaptations would not. The production of organismal diversity thus required that organisms be at once autonomous from, and sensitive to, the environment, another example of complexity.

Darwin visualized his viewpoint about the complexity of evolution with two metaphors, the phylogenetic tree and the tangled bank. The phylogenetic tree points to complexity arising from irreversible phenomena. By referring to species as “communities of descent” and placing them in a single “tree of life,” Darwin emphasized that the fundamental explanatory principle is shared history. The tangled bank, by contrast, points to complexity arising from biological associations that do not share the same temporal and spatial origins, producing reticulated (tangled) patterns of descent and association.

By the end of the 19th century, most biologists had adopted the view that evolution had occurred, but many were uncomfortable with Darwinism’s lack of simple general laws, and the discovery of the “Laws of Inheritance” by Mendel did little to quell this unrest. Bowler (1983) documented the eclipse of Darwinism during the period of 1890–1940. The most influential theories during that time were neo-Lamarckism and orthogenesis, both theories of simplicity. As a theory of environmental determinism, neo-Lamarckism placed causality solely in the nature of the conditions. Orthogeneticists, by contrast, asserted that the environment played little or no role in evolution, thereby placing causality solely in the nature of the organism. The eclipse of Darwinism is assumed to have ended with the emergence of neo-Darwinism in the 1930s. Neo-Darwinism focused on Darwin’s key mechanism, natural selection. Buoyed by advances in understanding the mechanisms of genetic inheritance

and some elegant experiments, this became the dominant theory for evolutionary biologists during the second half of the 20th century. Some founders fully expected neo-Darwinism to encompass the panorama of phenomena encompassed in Darwin's original formulation:

"...in every part of the whole, wonderful history of life, all the modes and all the factors of evolution are inextricably interwoven. The total process cannot be made simple, but it can be analyzed in part. It is not understood in all its appalling intricacy, but some understanding is in our grasp, and we may trust our own powers to obtain more."

—Simpson (1953: 393)

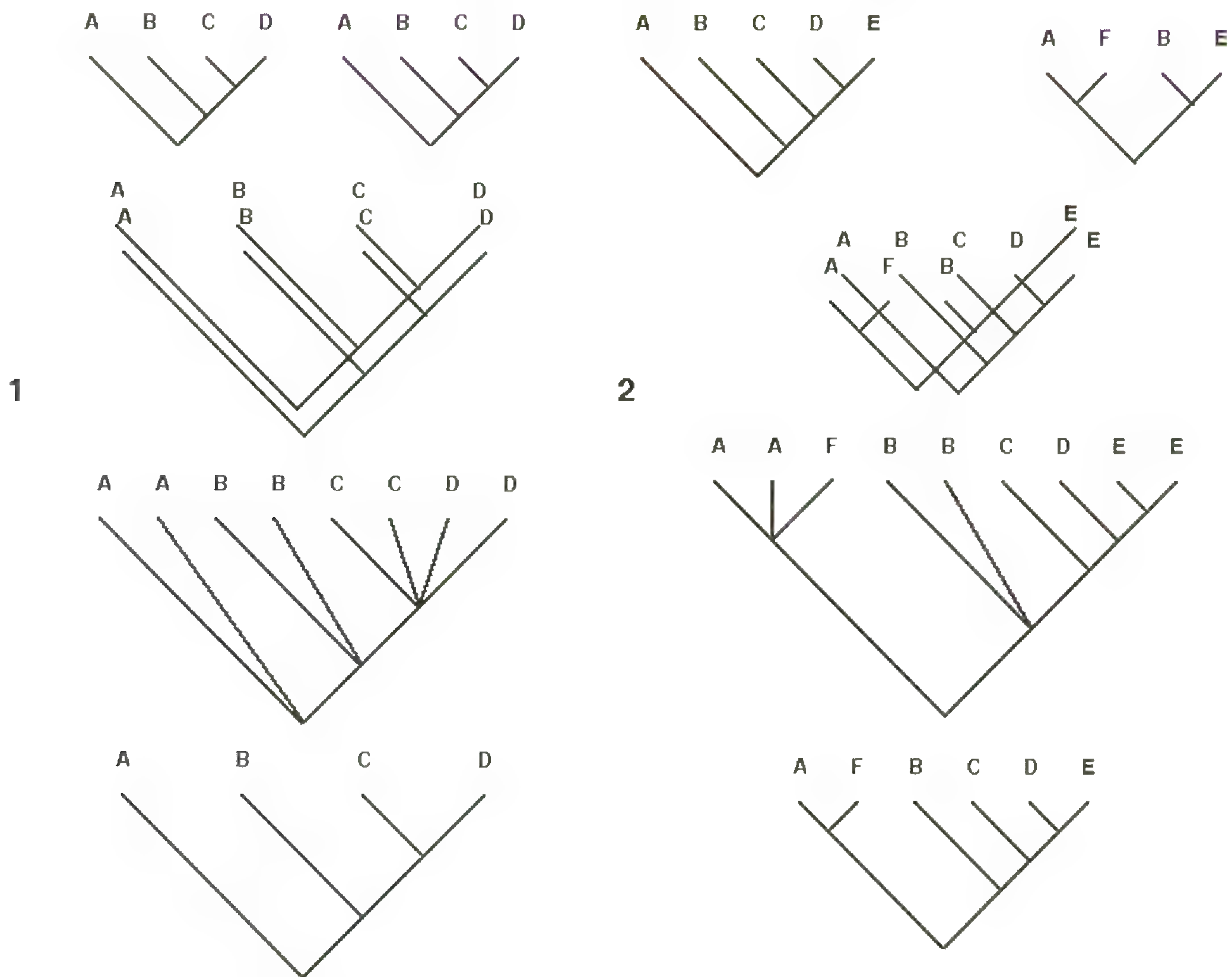
Sadly, this was not to be the case. As noted by Gould (1983) using the rubric "the hardening of the synthesis," during the latter third of the 20th century, neo-Darwinism maintained a narrow focus on population genetics, acknowledging other aspects of evolution but asserting that all of them were manifestations of population genetics and natural selection and, therefore, all appearances of complexity could be reduced to the sometimes computationally complex, yet ontologically simple, principles of population genetics. Like the neo-Lamarckians, neo-Darwinians focused their attention on the environmental conditions correlated with the survival of organisms, populations, and species. Thus, many explanations expounded by neo-Darwinians sound Lamarckian because the nature of the organism has been reduced to almost nothing. For many today, the nature of the organism is to be a blind watchmaker made up of selfish genes. Organisms exist, but they have no "nature" autonomous from their surroundings. This produces a theory that is parsimonious, yet lacking Darwin's panoramic view of biological diversity.

During the past 25 years, a growing number of scientists and philosophers have called for modifications of evolutionary theory (e.g., Eldredge, 1985, 1986, 1989; Brooks & Wiley, 1986, 1988; Maynard Smith & Szathmary, 1995; Brooks, 1998, 2000, 2001, 2002; Collier & Hooker, 1999; Brooks & McLennan, 2000 and references therein). Many of these calls seek to reintroduce Darwin's postulates about "the nature of the organism" and the "nature of the conditions" to develop analytical methods for integrating the tree of life and the tangled bank metaphors, and to establish fundamental guiding principles or null hypotheses for complex evolutionary phenomena occurring on different spatial and temporal scales.

The phylogenetics revolution that began in the late 1960s (Hennig, 1966) and became entrenched beginning in the 1980s (for a review, see Brooks &

McLennan, 2002) produced quantitative analytical methods for effectively documenting the complexity of the tree of life. Phylogeneticists discovered quickly that most phylogenetic tree complexity is manifested in a relatively small number of linearly organized historical correlations (which we consider to be the communities of descent) and, therefore, is (once again) computationally complex but ontologically simple. Methods developed for analyzing historical associations described by tangled bank complexity assumed that such associations could also be adequately described and explained using phylogenetic tree-based reasoning (e.g., Brooks, 1985). Tangled bank complexity, however, produces reticulated historical relationships in addition to the relatively simple linear ones. We now know that methods of analysis based on phylogenetic tree-based reasoning produce internally inconsistent results in direct proportion to the degree of tangled bank complexity in the data, an indication that those methods have oversimplified the data (van Veller et al., 1999, 2000, 2001, 2002, 2003; Dowling, 2002; Dowling et al., 2003). Wojcicki and Brooks (2004, 2005) developed an algorithm called phylogenetic analysis for comparing trees (PACT) that permits phylogenetic systematic principles to be used in describing tangled bank type complexity.

PACT is based on the assumption that tangled bank complexity is based on individual parts that each conform to relatively simple tree-like complexity which, when combined, may nonetheless produce complex patterns of both general and unique relationships, including the possibility of reticulated relationships. PACT uses three guiding principles: (1) Assumption 0 (Wiley, 1986, 1988a, b; Zandee & Roos, 1987): in order to not oversimplify, you must analyze all input data without modification, and your final result must be logically consistent with all input data; (2) the duplication rule (Brooks & McLennan, 2002): Assumption 0 is violated with analyses by methods of simplicity when the data result from reticulated processes. Oversimplification can be avoided (i.e., Assumption 0 can be satisfied) in such cases by duplicating entities with reticulated histories; and (3) epistemological parsimony (Brooks & McLennan, 2002): in order to find the maximum number of generalities supported by the data, entities should not be duplicated beyond necessity. Therefore, we should make only enough duplications to satisfy Assumption 0. Underscoring its connection with the ontology of complexity, note that Assumption 0 is more fundamental than parsimony. If evolutionary patterns were simple, Assumption 0 would never be violated, so we would never need to duplicate entities and would never need parsimony. What we are seeking is the most general explanation possible



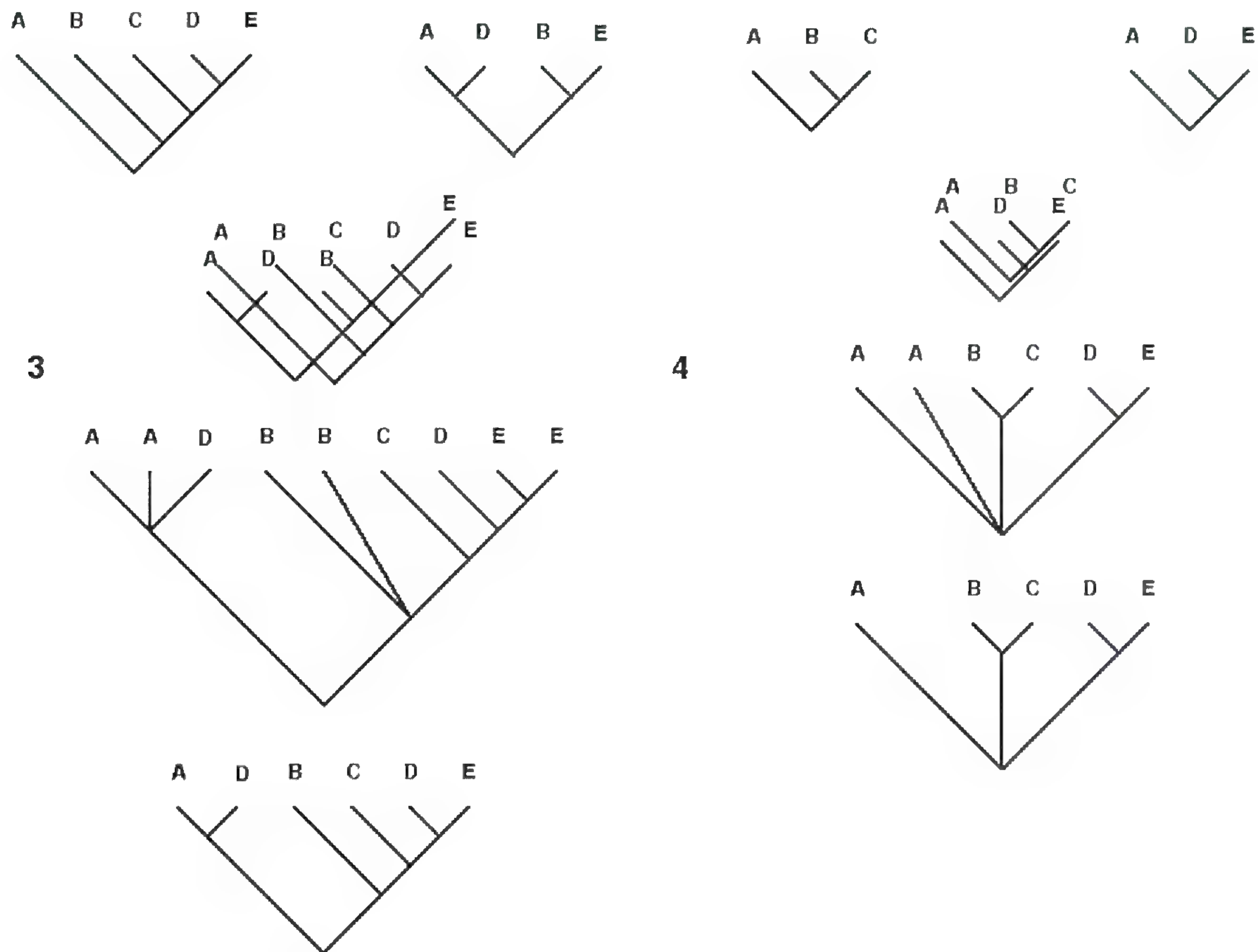
Figures 1 and 2. —1. PACT analysis for simple congruence. Top = two hierarchical patterns based on different data but pertaining to entities A, B, C, and D. Second from top = the two hierarchical patterns aligned by common elements. Third from top = both hierarchical patterns superimposed on each other. Bottom = PACT solution based on applying combination rule 1 ($Y + Y = Y$). —2. PACT analysis for two incongruent patterns. Top = two hierarchical patterns based on different data but pertaining to entities A, B, C, D, E, and F. Second from top = the two hierarchical patterns aligned by common elements. Third from top = both hierarchical patterns superimposed on each other. Bottom = PACT solution based on applying combination rule 2 ($Y + YN = YN$). $Y(A) + YN(AF) = YN(AF)$; $Y(E) + YN(ED) = YN(ED)$; $Y(B) + YN(C(DE)) = YN(B(C(DE)))$.

given the data (i.e., the ontology of complexity), not the simplest explanation we can imagine (i.e., the ontology of simplicity) (see van Veller & Brooks, 2001; Brooks & McLennan, 2002). What we have to guard against is the use of models and analytical methods that oversimplify the data in an effort to generalize. That is, simplicity is not always the most parsimonious depiction of the real world (van Veller & Brooks, 2001).

PACT utilizes any data that can be represented as hierarchical strings, ranging from gene trees and other character state trees to species phylogenies. Wojcicki and Brooks (2004, 2005) recognized that nested (historical) hierarchical patterns in nature are combinations of only six different classes of pattern modules. Analysis of a set of data is accomplished by selecting any one of the input data strings as the initial template. Next, the final analysis is built by sequentially aligning each additional data string with the template. Thereafter, combinations are made to satisfy Assumption 0 using four combination rules.

The combination rules are illustrated in Figures 1–6 and explained briefly in the figure legends; for more details, see Wojcicki and Brooks (2004, 2005). Some workers seem to have assumed that PACT is an algorithm for the method called Secondary BPA (Brooks & McLennan, 2002). A purported critique of PACT by Salvador Arias et al. (2008) addresses shortcomings of secondary BPA overcome by PACT that were specifically addressed by Wojcicki and Brooks (2004, 2005).

Elucidating complex evolutionary patterns using PACT comes with a cost of sorts. Explaining the complexity of a set of observations can only be achieved by abandoning a priori expectations of mechanism. For example, is the presence of a species in area F in Figure 2 an indication of a unique event involving the production of a novelty in one of the data strings, or is it an indication of a unique event involving loss from the other? PACT produces the same result regardless of the interpretation, so no assertions or assumptions about processes are em-



Figures 3 and 4. —3. PACT analysis for two incongruent patterns. Top = two hierarchical patterns based on different data but pertaining to entities A, B, C, D, and E. Second from top = the two hierarchical patterns aligned by common elements. Third from top = both hierarchical patterns superimposed on each other. Bottom = PACT solution based on applying combination rule 2 ($Y + YN = YN$). $Y(A) + YN(AD) = YN(AD)$; $Y(E) + YN(ED) = YN(ED)$; $Y(B) + N(C(DE)) = YN(B(C(DE)))$. Note that this example is precisely the same as the one shown in Figure 2, with the exception that in this case we find evidence of reticulated relationships. —4. PACT analysis for two congruent patterns. Top = two hierarchical patterns based on different data but pertaining to entities A, B, C, D, and E. Second from top = the two hierarchical patterns aligned by common elements. Third from top = both hierarchical patterns superimposed on each other. Bottom = PACT solution based on applying combination rule 3 ($YN + YN = YNN$). $Y(A)N(BC) + Y(A)N(DE) = YNN(A(BC)(DE))$. (One should not over-interpret the data.)

bedded in the algorithm. Similar reasoning can be applied to concepts such as lineage duplication, lineage sorting, and ancestral areas/hosts. All these concepts pertain to hypotheses of mechanism, and in order to maximize empirical robustness and avoid the possibility of circularity, all processes should be inferred following an analysis of data, not built into it.

It is for this reason that PACT makes combinations beginning with the youngest members of the data strings being compared. This avoids the assumptions, unintentionally embedded in all previous methods for historical biogeography and coevolution, that (1) all species associated today originated together at the same place and time; and (2) when there is ambiguity, we resolve the ambiguity beginning with the oldest (most basal) portions, amounting to an unacceptable assumption that the older the evolutionary event, the more accurate our information about it.

PACT analysis is, therefore, only the beginning of a study, not an end in itself. Explaining the result of a PACT analysis requires reference to guiding principles that are more comprehensive than those currently

in use. We illustrate this briefly for four different research programs in comparative evolutionary biology.

RESEARCH PROGRAMS

HISTORICAL BIOGEOGRAPHY

Biogeography is the basis for understanding the origin of species, the assemblage of communities, and the differential radiation of clades (Brooks & McLennan, 2002). In the latter third of the 20th century, historical biogeography produced two simple and elegant theories that were based on an ontology of simplicity rather than complexity.

The first of these was the equilibrium theory of island biogeography (ETIB) (MacArthur & Wilson, 1963, 1967). This theory is based on the view that dispersal from “source” areas to “islands” (actual or metaphorical), mediated by island size and distance, produces linear log-normal species-area relationships. Noise in the system or the effects of contingency comprise in situ speciation and extinction. From this,

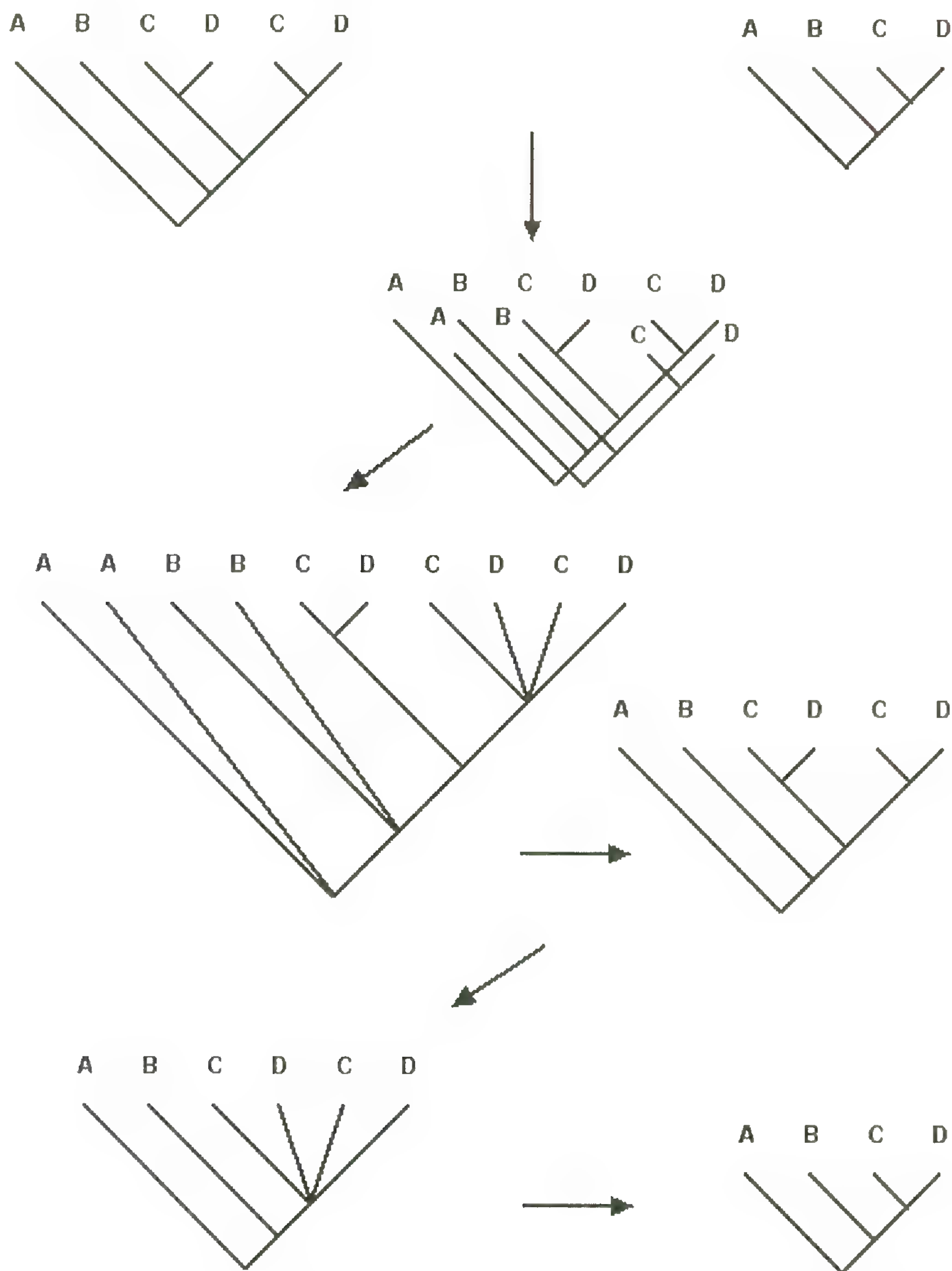


Figure 5. PACT analysis for lineage duplication (sympatric speciation). Top = two hierarchical patterns based on different data but pertaining to entities A, B, C, and D. Second from top = the two hierarchical patterns aligned by common elements. Third from top left = both hierarchical patterns superimposed on each other. Third from top right = PACT solution based on applying combination rule 1 ($Y + Y = Y$). Note that this leaves duplicate (CD) elements coming from a common node. Bottom left = two (CD) groups coming from a common node superimposed on each other. Bottom right = PACT solution based on applying combination rule 1 ($Y + Y = Y$).

one infers that data that conflict with the expected pattern (the “law”) are the result of historical contingencies, and it is therefore permissible to remove or modify them. As a result, island biogeographers are admonished to study small, young islands in order to minimize the potential for such historical contingencies that cloud our ability to see the true (and simple) pattern.

The second theory of simplicity was the maximum vicariance hypothesis (MaxVic), also known as vicariance biogeography or cladistic biogeography (Humphries & Parenti, 1999). MaxVic is based on the theory that in situ speciation and extinction produce simple area cladograms in which each area occupies a unique position. Noise in the system or the effects of

contingency result from dispersal. From this, one infers that data that conflict with a single area cladogram in which each area appears once (the “law”) are the result of historical contingencies, and it is therefore permissible to remove or modify them. As a result, cladistic biogeographers developed Assumptions 1 and 2 to remove or modify (“reconcile”) incongruent data with a single simple area cladogram in which each area occupies a unique position.

One persistent concern about both these paradigms has been this: if it is necessary to remove and modify data and to restrict one’s scope of analysis, just how general and powerful are the explanations produced? A closer comparison of the two paradigms reveals another interesting feature: they are complementary

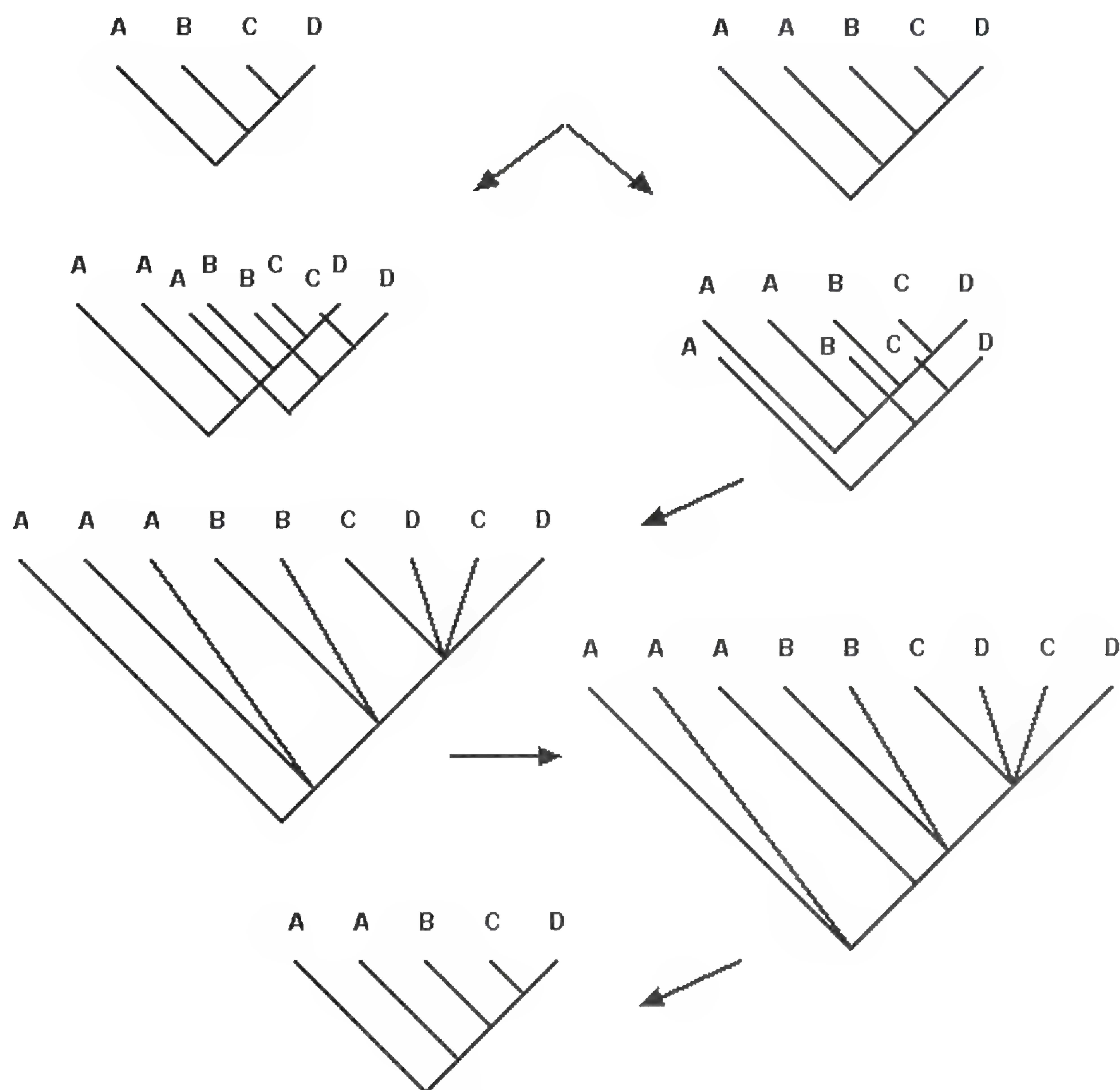


Figure 6. PACT analysis of two partly congruent patterns. Top = two hierarchical patterns based on different data but pertaining to entities A, B, C, and D. Second from top = the two hierarchical patterns aligned by common elements. Third from top = both hierarchical patterns superimposed on each other. Bottom = PACT solution based on applying combination rule 4 ($Y(Y- + Y = (Y(Y-, Y(A) (Y- (A- + Y(A-) = Y(A) (Y(A-$. This exemplar is also known as the “problem of paralogy.”

theories, each one excluding the other’s domain of explanation. What is missing from ETIB are assessments of the geographic origin of species, even though this is what distinguishes a “source” from an “island.” What is missing in MaxVic are assessments of post-speciation movements, and yet this is how ancestral species become widespread enough to be affected by vicariance. If each of these theories describes something valid and each one excludes the other’s explanatory domain, perhaps the problem lies in the adoption of ontological parsimony on the part of advocates of both theories of biogeography.

A “new” guiding principle: The taxon pulse. At nearly the same time the maximum vicariance paradigm emerged, Erwin (1979, 1981) proposed the taxon pulse hypothesis as a model incorporating both dispersal and vicariance. Erwin’s model stemmed from an idea proposed by Darlington (1943), later named the “taxon cycle” by Wilson (1959, 1961).

Taxon pulse and taxon cycle models both assume that species and their adaptations arise in “centers of diversification,” and that distributional ranges of taxa periodically fluctuate around a more stable, continuously occupied center. This general biotic dispersal may be interrupted by the formation of barriers, producing episodes of vicariant speciation. Breakdown of those barriers produces new episodes of biotic expansion, setting the stage for yet more episodes of vicariance. Taxon cycles occur over relatively short periods of time (“ecological time”) and involve species that disperse actively and colonize new areas during expansion episodes, then contract their ranges during periods of habitat contraction without producing new species. Taxon pulses, by contrast, occur over relatively long periods of time (“evolutionary time”) and are characterized by dispersal along a broad front during expansion into suitable habitat when previous barriers break down. During this expansion phase, different species within

a biota encounter additional geographic heterogeneity, including range contractions. Such heterogeneity may (1) stop the expansion of some species, resulting in species of restricted distributions; (2) affect only the rate of expansion for some species, producing widespread species; or (3) act as barriers to dispersal of sufficient magnitude to produce new species as a result of peripheral isolates speciation. Geological evolution, operating on longer time scales than biological evolution, may also produce barriers, resulting in episodes of vicariant speciation affecting members of these same biotas.

Despite the existence of an alternative to MaxVic and despite concerns that exemplar taxa were being carefully selected to show a preponderance of vicariance (Simberloff et al., 1980; Simberloff, 1987), vicariance has become the default explanation for any observation of allopatry. And yet, MaxVic has always been deficient because it neglects the issue of how ancestral species of many clades become widespread enough to be affected by vicariant events. If vicariance affects many members of ancestral biotas in the same way, it seems reasonable to assume that at some point in the past the members of the biota expanded their geographic ranges to such an extent that they could be affected by the subsequent vicariance event. Advocates of vicariance biogeography have acknowledged that this must happen. Wiley (1981) noted that some circumstances, such as colonization of islands, might produce general distribution patterns based on dispersal rather than vicariance, and Endler (1982) suggested that such correlated dispersal patterns might be common. In practice, however, historical biogeographers have simply assumed that such dispersal does not produce general patterns, so it is permissible to invoke dispersal only to explain departures from the general pattern, which is always explained as the result of vicariance (Wiley, 1986, 1988a, b; Brooks & McLennan, 1991, 2002; Humphries & Parenti, 1999).

Taxon pulse-driven biotic diversification differs from vicariance-driven biotic diversification in three important ways. First, because diversification is driven by biotic expansion, we expect to find general patterns associated with dispersal, not just with vicariance. General patterns resulting from biotic expansion occur when barriers to dispersal, especially the large-scale ones leading to vicariance, break down. Second, episodes of biotic expansion, even those involving large areas, will inevitably lead to reticulated historical relationships among areas and biotas within areas of endemism comprising species of different ages derived from different sources. Third, the absence of particular clades in particular areas is more parsimoniously explained as a lack of participation in that particular

expansion episode by a particular clade, rather than dispersal with extinction. Taxon pulses are also historically contingent, meaning that at any given time, different clades comprising a complex biota may form a mosaic of area relationships.

Recent studies have shown extensive reticulated area relationships even for data sets carefully chosen to emphasize vicariance (e.g., Brooks & McLennan, 2001; Green et al., 2002; McLennan & Brooks, 2002; Spironello & Brooks, 2003; Bouchard et al., 2004; Brooks & Ferrao, 2005; Brooks & Folinsbee, 2005; Halas et al., 2005; Folinsbee & Brooks, 2007).

A new research program for historical biogeographic analyses. Following Halas et al. (2005), a new research program based on an ontology of complexity for historical biogeographical analyses comprises two steps:

Step 1—Producing a general area cladogram.

Convert all phylogenetic trees into taxon-area cladograms by replacing the names of the species with the areas they inhabit, then combine them using PACT, producing a general area cladogram (GAC). Figure 7 is the GAC resulting from PACT analysis of elephants, hyenas, and hominoids since the Miocene (Brooks & Folinsbee, 2005; Folinsbee & Brooks, 2007).

Step 2—Inferring biogeographic history from the GAC.

This is a two-part process. First, we must distinguish general from unique nodes. Unique nodes are produced by an evolutionary event affecting only a single clade. At present, all nodes associated with evolutionary events affecting more than one clade are termed general nodes. Clearly, a general node involving two of 100 clades likely is not the same as one involving 75 of 100 clades. At the moment, however, we do not have enough empirical data or any models to allow us to be more fine-grained in our distinctions among general nodes, but this is clearly an area of interest for future research. Second, we must interpret the general nodes. Until recently, historical biogeographers believed that all general nodes in an area cladogram should be ascribed to vicariance. Lieberman (2000, 2003a, b), however, noted that episodes of geodispersal (geographic areas fusing rather than splitting) could produce general nodes resulting from dispersal. In a similar vein, the taxon pulse hypothesis asserts that episodes of general biotic expansion occurring between vicariance (or, more generally, isolation) events set the stage for future isolation events. Lieberman (2000, 2003a, b) proposed a protocol for distinguishing general nodes in area cladograms due to isolation from those caused by geodispersal or biotic expansion. The protocol is based on phylogenetic character optimization (Fitch optimization), using areas inhabited as the characters.

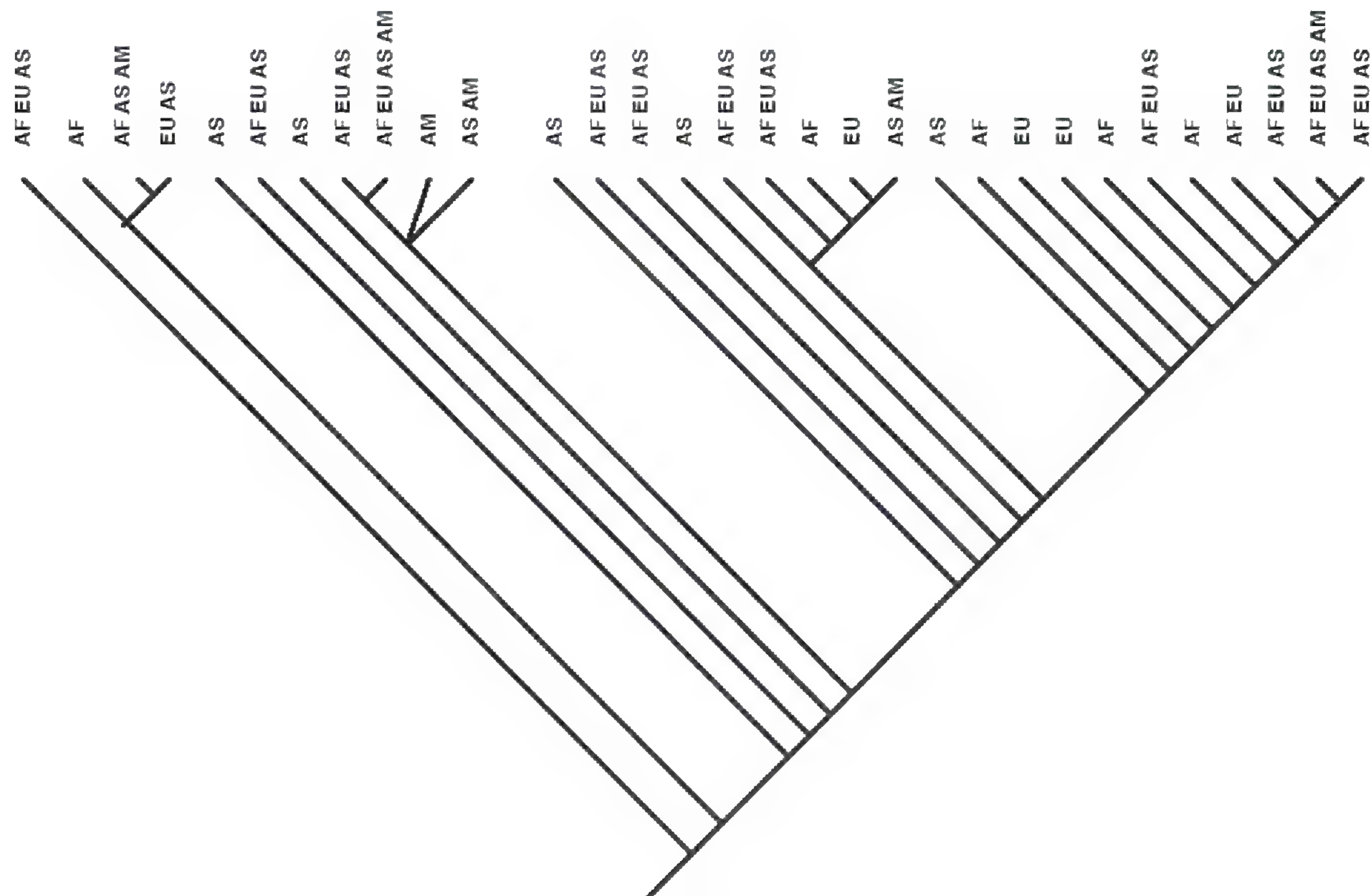


Figure 7. PACT analysis of elephants, hyenas, and hominoids since the Miocene (Brooks & Folinsbee, 2005; Folinsbee & Brooks, 2007). General area cladogram. AF = Africa, AM = Americas, AS = Asia, EU = Europe.

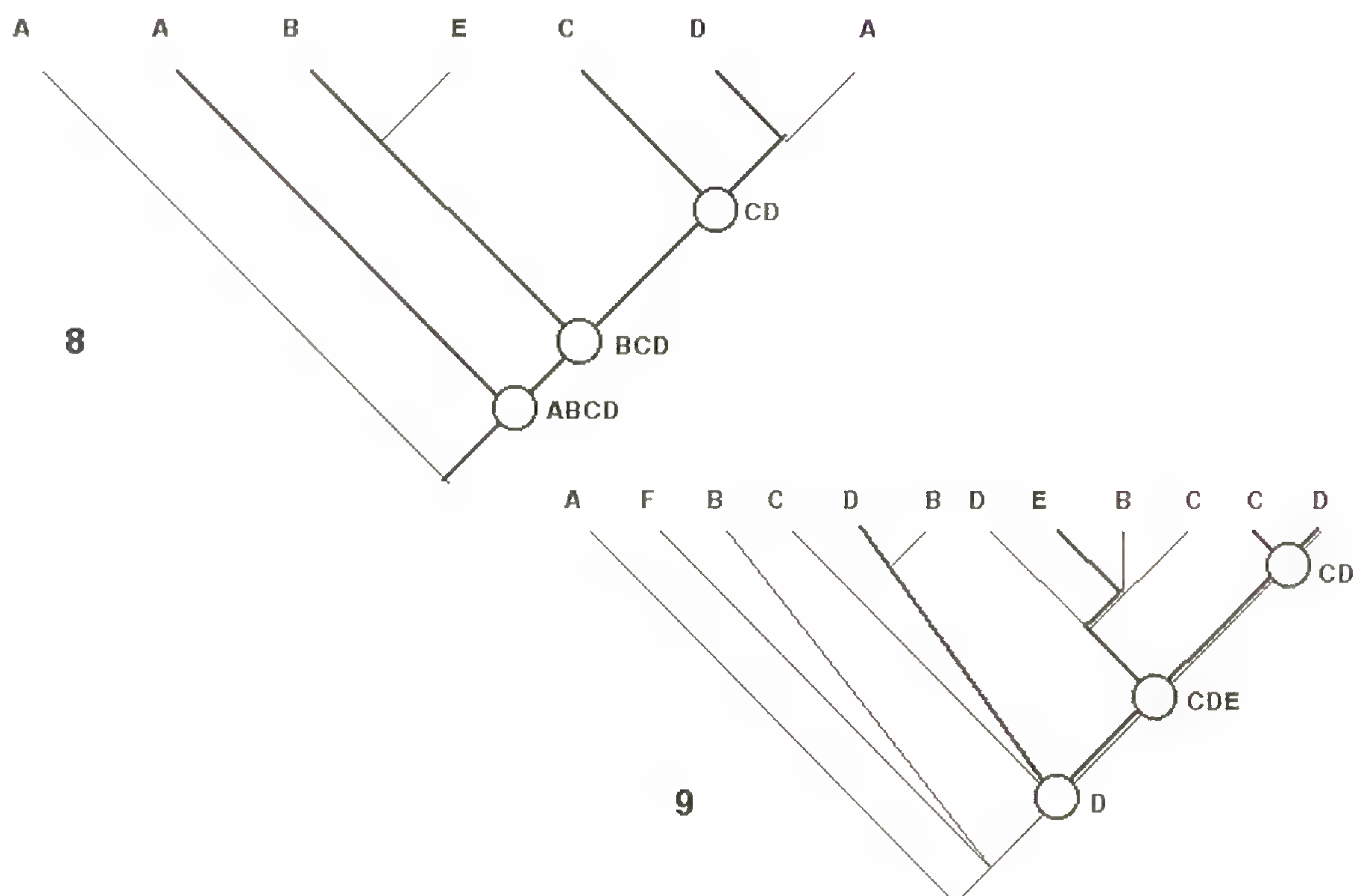
General nodes associated with vicariance or other forms of isolation exhibit decreasing numbers of areas occupied relative to the next oldest node (Fig. 8), whereas general nodes associated with biotic expansion (breakdown of a barrier) are associated with increasing numbers of areas occupied relative to the next oldest node (Fig. 9).

Figure 10 summarizes the nodal analysis for the GAC shown in Figure 7. The alternation of isolation (V, or vicariance) nodes with biotic expansion (BE) and within-area (W) differentiation nodes produces the characteristic signature of a taxon pulse radiation. The relatively large number of general nodes (22 out of 28, or 79%) suggests that all three clades participated in the same general taxon pulse scenario, summarized in Figure 11.

Complexity studies in biogeography: Integrating historical and ecological biogeography. That larger islands have greater species richness than smaller islands, and that “islands” need not be oceanic because species richness increases with any increased sample of area, have long been recognized. ETIB (MacArthur & Wilson, 1963, 1967) predicts a linear log-normal function relationship between species richness and the size of an island resulting from a dynamic balance between immigration, that is, colonization from a source area, and extinction. The extinction rate is assumed to increase with the number of species present on any island, so that small areas with higher species richness than expected have a

higher extinction rate than immigration rate and are not yet in equilibrium. MacArthur and Wilson (1963) acknowledged that “local speciation” (in situ speciation) would confound the species-area relationship, but suggested that for most cases it was probably safe to omit in situ speciation from the model as its effect on the species-area relation is “probably significant only in the oldest, largest, and most isolated islands” (1963: 380). Recent discussions, however, have suggested that such historical phenomena require closer investigation (e.g., Heaney, 2000; Whittaker, 2000).

Halas et al. (2005) performed the first direct analysis of the phylogenetic impact on the species-area relationship using the extensive data set presented by Marshall and Liebherr (2000), representing 33 clades of insects, vertebrates, and flowering plants occurring throughout Mexico and parts of Central America. Nodal analysis of their GAC permitted them to classify each species in each area as a colonizer or a resident. Correlating species richness and area size for the 33 clades and nine areas in Mexico and Central America (for details, see Halas et al., 2005) produced a low correlation coefficient for the species-area curve ($r^2 = 0.47$), due primarily to two relatively small areas (the Transmexican Volcanic Belt and the Sierra Madre del Sur) containing unusually large numbers of species. Correlating extinction events and species richness for this data set produced a high correlation coefficient ($r^2 = 0.75$), indicating strong support for this prediction of the



Figures 8 and 9. —8. Lieberman's nodal analysis 1. Thick branches indicate general nodes; thin branches represent elements of the GAC produced by only one clade. Progressive decrease in the number of areas associated with nodal values indicates that all general nodes are isolation nodes. This corresponds to either maximum vicariance or Hennig's progression rule. —9. Lieberman's nodal analysis 2. Thick branches indicate general nodes; thin branches represent elements of the GAC produced by only one clade. Increase in the number of areas associated with nodal values from node D to node CDE indicates an episode of biotic expansion or general dispersal; decrease in the number of areas associated with nodal values from CDE to CD indicates an episode of isolation, corresponding to either vicariance or Hennig's progression rule. Alternation of biotic expansion and isolation nodes is characteristic of taxon pulse radiations.

ETIB. Following MacArthur and Wilson's lead, Losos and Schluter (2000) reported that, for anoline lizards on large Caribbean islands, inferred extinction rates were low and in situ speciation was a more important source of species richness than colonization. They predicted that effects similar to those they observed on the largest islands should be found on continental islands. Halas et al. (2005) corroborated those predictions: in situ speciation correlates better with area size than does colonization, and inferred extinction rates are low. The protocol produces inferences of only 19 extinction events involving 16 of the 33 clades, because direct parsimonious inferences of extinction can be made only for episodes of vicariance. If extinction rates are the same following biotic expansion events, which account for 60% of the general nodes, the number of inferred extinctions increases from 19 to 48, but does not change the correlations with area size.

These data provided additional insight into the evolutionary relationship between colonization and in situ speciation. There is a very poor correlation between colonization and in situ speciation ($r^2 = 0.02$), indicating that these phenomena are relatively independent of each other. We suggest that the reason historical effects on large islands confound the species-area relationship is the subsequent dispersal

of some species produced in situ to other islands, so sources become islands and islands become sources on evolutionary time scales. All nine areas studied by Halas et al. (2005) have served as both sources and islands at different times and to different degrees. Halas et al. (2005) identified six fundamental episodes of vicariance and nine of biotic expansion as the main determinants of the biogeographic patterns observed in the data set. The Transmexican Volcanic Belt has acted as an "island," receiving species by colonization, for all nine biotic expansion events, and the Sierra Madre del Sur has acted as a dispersal source for three of the episodes of biotic expansion. At the same time, both areas were involved in five of the six episodes of vicariance. This explains why these relatively small areas are disproportionately species rich, without any evidence of an accompanying high extinction rate.

The analysis by Halas et al. (2005) indicates that MacArthur and Wilson (1963, 1967) were correct—in situ speciation confounds the species-area relationship. This occurs when the same areas function as islands and as sources at different times. How can this happen? The answer is niche diversification associated with speciation, producing island species that can colonize the source area(s). Niche diversification most often occurs at a much slower rate than

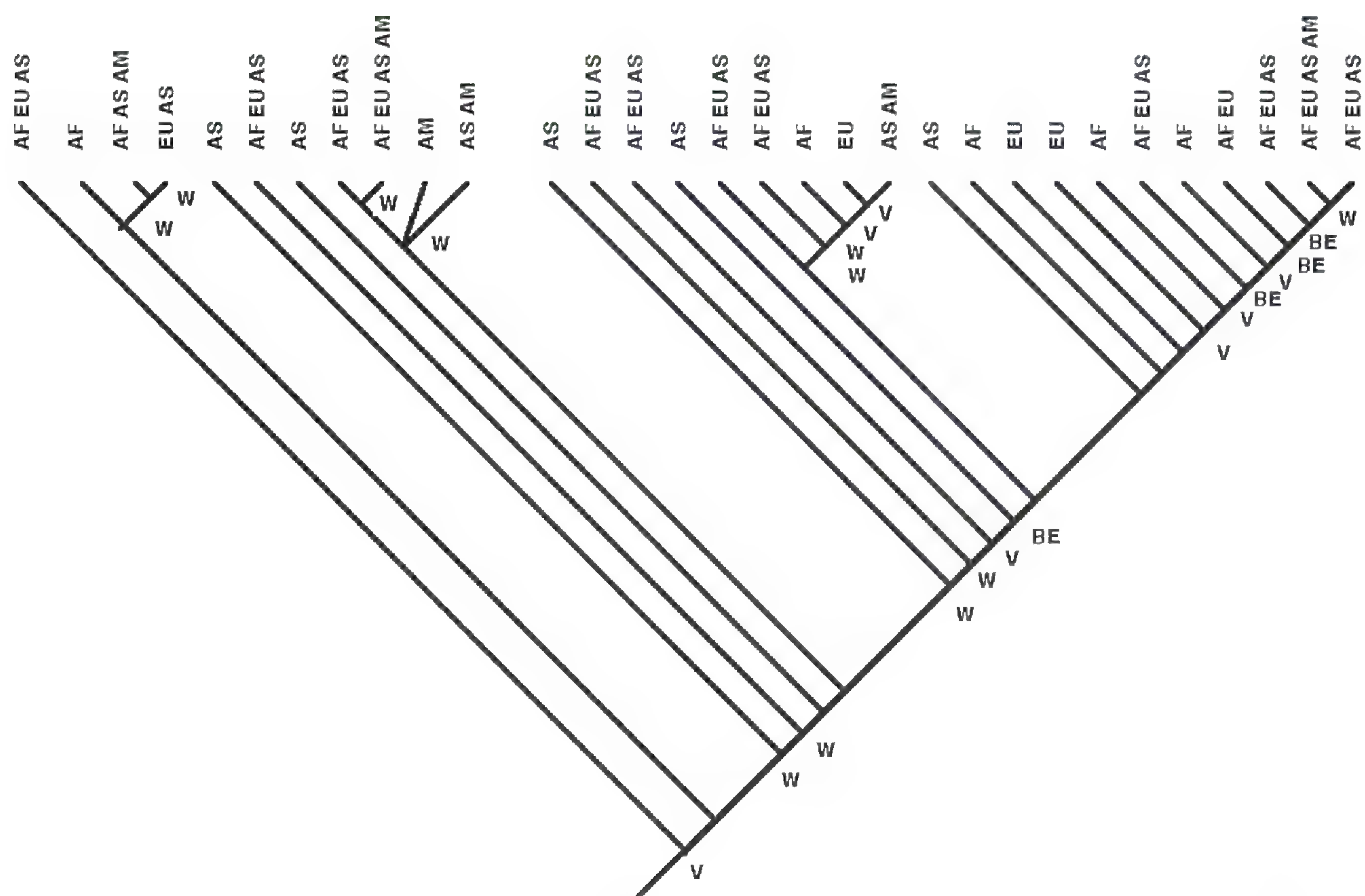


Figure 10. PACT analysis of elephants, hyenas, and hominoids since the Miocene (Brooks & Folinsbee, 2005; Folinsbee & Brooks, 2007). General node analysis. AF = Africa, AM = Americas, AS = Asia, BE = biotic expansion, EU = Europe, V = vicariance, W = within area. Alternation of biotic expansion/within-area and vicariance (isolation) nodes is characteristic of taxon pulse radiations.

speciation (Brooks & McLennan, 2002), but the biogeographic “role reversal” resulting from niche diversification need not occur often in order to affect the equilibrium number of species in a given area. In addition, Cracraft (1985) seems to have been correct in asserting that speciation and extinction are under different rate controls. In the study by Halas et al. (2005), extinction rates correlate well with area size, but speciation rates do not. How can this happen?

According to Cracraft (1985), speciation is largely controlled by geography and extinction by environmental harshness, associated with adaptation and niche diversification. That is, extinction is a failure to adapt, not a failure to speciate (Brooks & McLennan, 2002). Occasional episodes of niche diversification plus episodic changes in environmental harshness are the primary source of the “pulses” in the taxon pulse hypothesis. The study by Halas et al. (2005) suggests

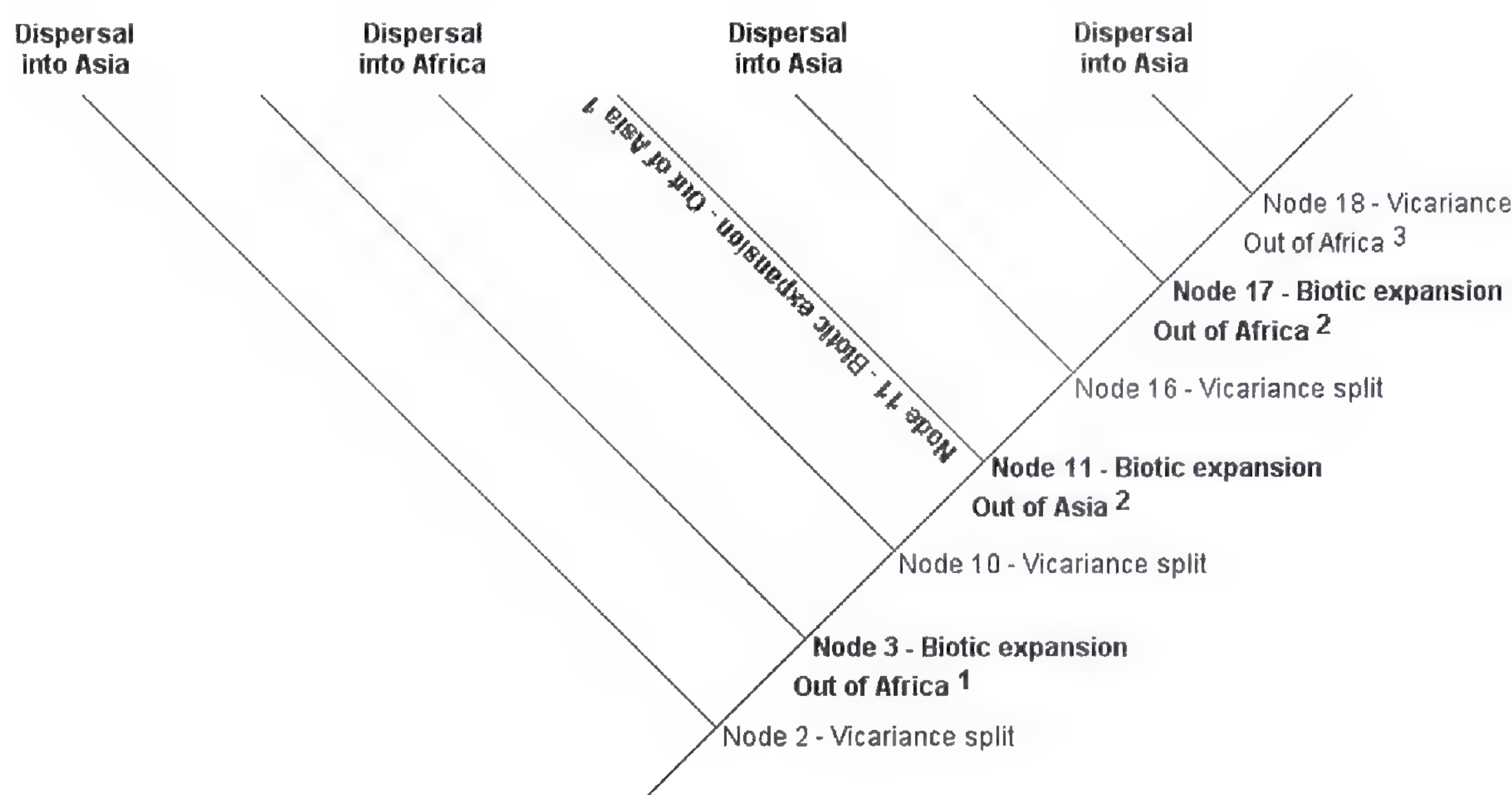


Figure 11. PACT analysis of elephants, hyenas, and hominoids since the Miocene (Brooks & Folinsbee, 2005; Folinsbee & Brooks, 2007). Pictorial representation of the taxon pulse scenario that best explains the data.

that Losos and Schluter (2000) were not correct in one respect: in situ speciation and colonization are not necessarily causally linked, such that high rates of one need not cause low rates in the other. If speciation and extinction are under different rate controls, the appearance of fit to their simple deterministic model was likely the result of using only a single clade for their analysis.

The combination of speciation and niche diversification makes the equilibrium number of species a moving target in evolutionary time; that is, evolution is a non-equilibrium phenomenon (Brooks & Wiley, 1988). This complements recent calls by, e.g., Heaney (2000) and Whittaker (2000) for modifications of the ETIB to incorporate complex patterns of immigration, extinction, and diversification occurring on various spatial scales and on both ecological and evolutionary time scales.

COEVOLUTION

Fahrenholz (1913) focused entirely on the hosts when he postulated that the occurrence of related blood-sucking lice on different primates demonstrated that the catarrhines were more closely related to hominoids than to any other primates. By the late 1930s, this orthogenetic perspective had produced an integrated view of coevolution: parasites are highly host specific, so they coevolve with their hosts, and because they coevolve with their hosts, they become highly host specific. Because host specificity was the *cause* of coevolution rather than a function of the ecological interaction between lineages, any conflicting or inconsistent observations were considered to be erroneous or irrelevant because they failed to conform to the orthogenetic view of coevolution. Vestiges of orthogenetic thinking persist today; this is especially true for the assumption that hosts and parasites “ought” to have congruent phylogenies, embodied in the maximum cospeciation school of research (Page, 2002).

In contrast with the orthogenetic view of coevolution, Kellogg (1896, 1913) cast his observations about host-parasite relationships in a Darwinian framework. His studies of birds and their biting lice suggested that while some host-parasite systems might show strong phylogenetic associations, there were substantial cases of what he termed “straggling” or “host-switching.” This perspective was adopted primarily by researchers interested in studying the interactions between plants and phytophagous insects (e.g., Verschaffelt, 1910; Brues, 1920, 1924). Because those associations often showed no clear phylogenetic component with respect to host species (though they often were extremely specific), researchers in this

tradition focused on discovering the ecological ties between organisms, particularly the cues insects used to locate their host plants. The modern version of this second perspective on coevolution emerged in the 1960s. Following a mathematical model proposed by Mode (1958), Ehrlich and Raven (1964) hypothesized that the evolutionary diversification of plants and insects had been fueled by complex coevolutionary interactions involving mutual modification. As suggested initially by Kellogg, such coevolutionary dynamics might have a general phylogenetic context, but need not parallel the evolutionary history of the specific taxa involved. The distribution of insects among plants followed the evolution of host resources and the evolution of insects’ abilities to utilize those resources, rather than the evolution of host species themselves.

Janzen (1968, 1973a, b, 1980, 1981, 1983, 1985a, b) argued that the appearance of tight coevolutionary associations at any single locality could be misleading. No matter where a given species evolved in the first place, its inherited functional abilities may allow it to survive in a variety of places under a variety of conditions through arbitrary amounts of time. In other words, species and their phylogenetically conservative traits may disperse through time and space. He called this interaction between the past history of the species and their present-day associations “ecological fitting” (Janzen, 1985b).

All perspectives discussed above agree that parasites are resource specialists. They disagree fundamentally on the question of host switching. Can you be a resource specialist and still host switch readily? Brooks and McLennan (2002) discussed a number of ways in which this could happen. Beginning with the parasite, a species might be a resource specialist but also might share that specialist trait with one or more close relatives. That is, specialization on a particular resource can be a plesiomorphic characteristic of a clade of parasites. As a result, a given host species occurring in more than one area might be inhabited by two different species of parasites, related to each other but not necessarily as sister species. Such persistent plesiomorphic traits also might be co-opted to perform novel functions. Trouvé et al. (1998), for example, reported that life history traits of parasitic flatworms do not differ from life history traits of their free-living relatives, indicating that these species do not have a “parasitic mode of life” but rather a “platyhelminth mode of life” that has been co-opted to function in the context of parasitizing vertebrates. Persistent ancestral traits also might be “anachronisms” (Janzen & Martin, 1982), traits that evolved in a coevolutionary context that no longer exists. For example, a number

of members of the Neodermata (the clade of platyhelminths including trematodes, monogeneans, tapeworms, and their relatives) have guts, even though they also have a highly absorptive integument (the neodermis). In the case of the tapeworms, having a neodermis facilitated survival even when the gut was lost. The persistent intestine in other neodermatans is an anachronism, like the vermiform appendix in humans.

Alternatively and from the hosts' perspective, the resources themselves might be very specific and yet still taxonomically and geographically widespread. In such cases, a given parasite might inhabit more than one species of host, the hosts not being sister species, as a result of host switching. Two major categories of parasite specialization are preferred site of infection (Adamson & Caira, 1994; Brooks & McLennan, 1993), as well as their transmission dynamics (Brooks & McLennan, 1993; Hoberg & Adams, 2000; Hoberg et al., 2000).^{*} Phylogenetic conservatism in parasite biology, coupled with phylogenetic conservatism in host biology, would create a very large arena for host switching, even without the evolution of novel capabilities for host utilization.

If historical conservatism in the type of specialization, rather than specialization per se, determines the ease or difficulty of host switching, the extent of such historical conservatism should dictate our expectations about the frequency of occurrence of host switching. Two ubiquitous findings from the historical ecology revolution of the past 20 years are that all aspects of evolution, including ecology and behavior, are phylogenetically conservative, and that host switching is a regular feature of coevolutionary history (Brooks & McLennan, 1991, 1993, 2002). Moving phylogenetic studies of coevolution into the age of complexity thus requires the use of methods of analysis based on the recognition that host switching is common, not rare, and that the fundamental job of the investigator is to explain the switches that have occurred and, where possible, to anticipate switches that could occur rapidly during periods of substantial climate change.

A new guiding principle: The oscillation hypothesis. In comparative studies of coevolution, host shifts, including increasing host range and speciation by host switching, are the analog of dispersal in historical biogeography. Complexity studies of coevolution should, thus, be based on an analog of the taxon pulse hypothesis. Phylogenesis in taxa exhibiting "close and evident" ecological associations with host taxa should be characterized by episodes of increasing host range followed by episodes of isolation on particular hosts. Janz and Nylin (2007) have recently proposed such a model, which they call the "oscillation hypothesis" (Janz et al., 2006; Nylin &

Janz, 2007), and provided empirical evidence supporting it based on studies of various lepidopteran taxa and their host plants.

A new research program for coevolution studies. Within the realm of an ontologically complex research program, comparative studies of host-parasite associations use the following two steps:

Step 1—Producing a general host cladogram. Convert all parasite phylogenetic trees into parasite-host cladograms. This is accomplished by replacing the names of the parasite species with the hosts they inhabit. Then combine them using PACT—this produces a general host cladogram (GHC), a map of the host context of parasite speciation. We illustrate this approach using the pinworms (*Enterobius* Leach [1853]) and hookworms (*Oesophagostomum* [*Conoweb-eria*] Ihle [1922]) inhabiting the great apes including humans (for details with references, see Brooks & McLennan, 2003; Brooks & Ferrao, 2005). The GHC is shown in Figure 12.

Step 2—Assessing the host switches. The heavy branches on the host cladogram (Fig. 12) indicate instances of congruence between host and parasite phylogeny. This includes one case (*Enterobius gregorii* Hugot [1983] and *E. vermicularis* [Linnaeus, 1758] Leach [1853], both inhabiting humans) in which a parasite clade speciated but the host did not, and another (*Oesophagostomum stephanostomum* Stossich, 1904) in which the host clade speciated but the parasite did not. Despite the substantial amount of cospeciation between the parasites, which also happens to be congruent with the host phylogeny, the PACT analysis suggests that about 30% of the observed host associations are due to host switching. Only one represents a switch to a non-primate host (rodents). The remainder, all of which involve primate to primate switches, include acquiring parasites that are more or less the same age (three cases), that are younger (one case), and that are older (three cases) than the hosts.

Complexity studies in coevolution: Integrating biogeography and coevolution. We expect the evolution of increased host range to occur over relatively long time periods (Janz et al., 2006; Janz & Nylin, 2007; Nylin & Janz, 2007). This has led evolutionary biologists to predict that while generalists may be at an evolutionary disadvantage relative to specialists with respect to particular host species in the short term, they are at an advantage in the long term. Generalists have this advantage over specialists because of their ability to utilize multiple hosts, which increases their chances of survival during episodes of major climate change and

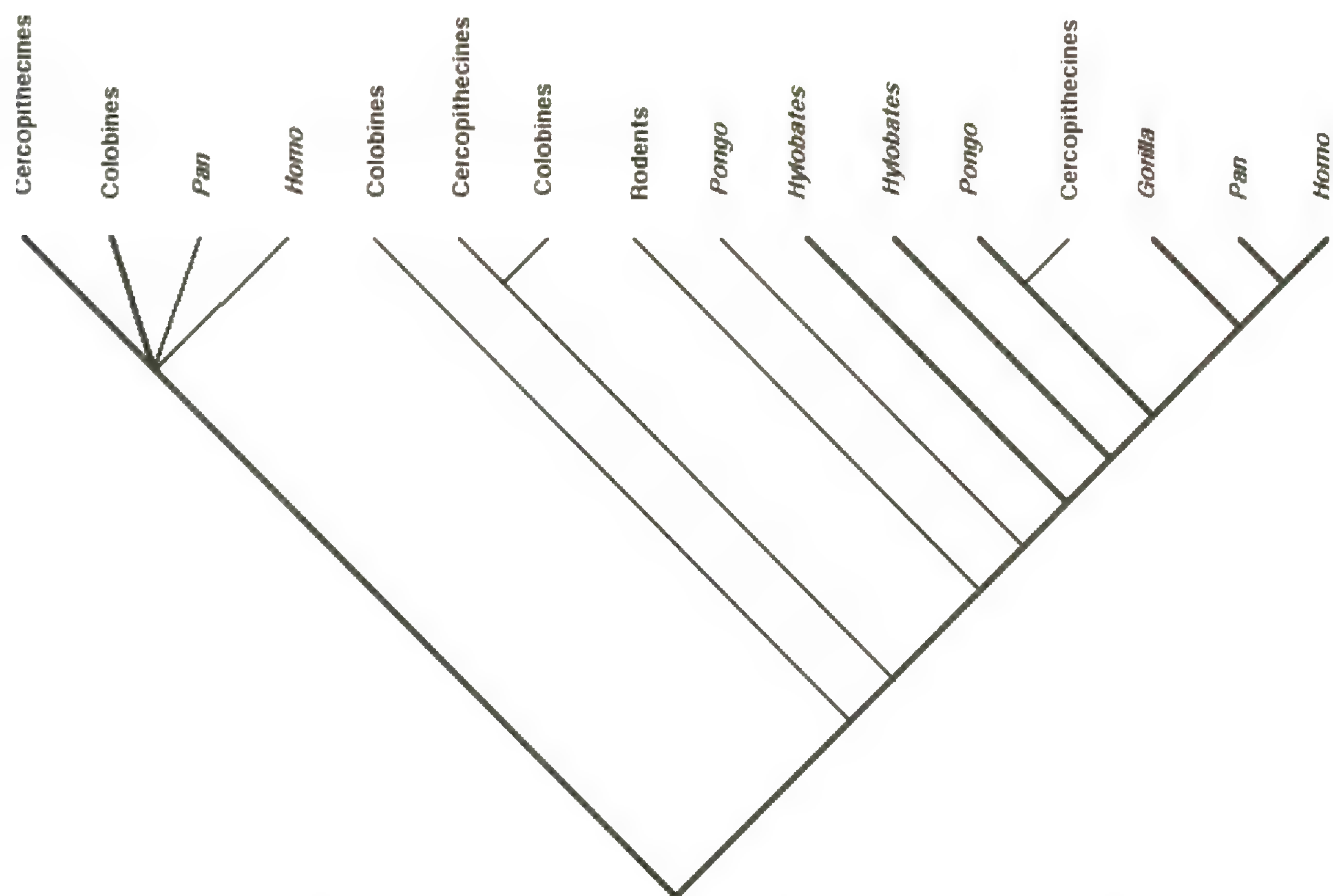


Figure 12. General host cladogram produced by PACT, indicating host context of speciation events implied by phylogenies for *Enterobius* Leach (1853) and *Oesophagostomum* (*Conoweberia*) Ihle (1922). Heavy lines indicate associations congruent with host phylogeny; thin lines indicate host-switching events.

associated changes in species composition and trophic structure in affected ecosystems. Brooks and McLennan (2002) suggested that ecological fitting, in the form of phylogenetically conservative host utilization capabilities, produces a more complex situation. In addition to true generalists, actively using different host species representing different resources, and true specialists, capable of surviving in association with only a single species of hosts, there are what Brooks and McLennan (2002) called “faux generalists” and “faux specialists.” Faux generalists are resource specialists whose resource is widespread among host species (synapomorphic, symplesiomorphic, or homoplasious). At any given place and time, only a restricted subset (often only a single species) of all potential host species are available to the specialist, but over a large geographic range, the specialist may be associated with many hosts. By having a large class of potential hosts, this class of specialists may have the same long-term evolutionary advantages as true generalists without giving up their short-term advantages as resource specialists. Faux specialists, by contrast, are resource generalists who, at any given place and time, are excluded from some suitable hosts by specialists on those hosts. By participating in associations with a small number of hosts, such generalists may have many of the same short-term benefits as true specialists, without losing the long-term advantages of being a resource generalist.

The most important aspect of this recognition by Brooks and McLennan (2002) is that true generalists, faux generalists, and faux specialists can all switch hosts rapidly during episodes of climate change. This has significant implications for public, livestock, and wildlife health specialists (Brooks & Hoberg, 2006, 2007a, b). One critical implication of the maximum cospeciation perspective is the expectation that progressive coevolutionary specialization mitigates host switches. This means that coevolution should provide its own safeguards against emerging infectious diseases (EIDs). We would thus expect EIDs to be rare events evolutionarily, and the current crisis would be seen as something different in kind rather than in degree from the evolutionary background of coevolution. Under the Darwinian view, by contrast, EIDs are not rare events but are common evolutionary accidents waiting to happen. In that case, the current EID crisis is different only in degree, not in kind, from coevolutionary dynamics occurring prior to the present. The potential for host switching is therefore proportional to the number of susceptible hosts that are not “apparent” to a parasite at any given time. The apparency arena, although phylogenetically conservative, is not static. Environmental changes in the place of origin for any parasite, leading to altered trophic interactions or to geographic dispersal, increase the chances that the parasite will come into contact with additional susceptible hosts. Each time this happens,

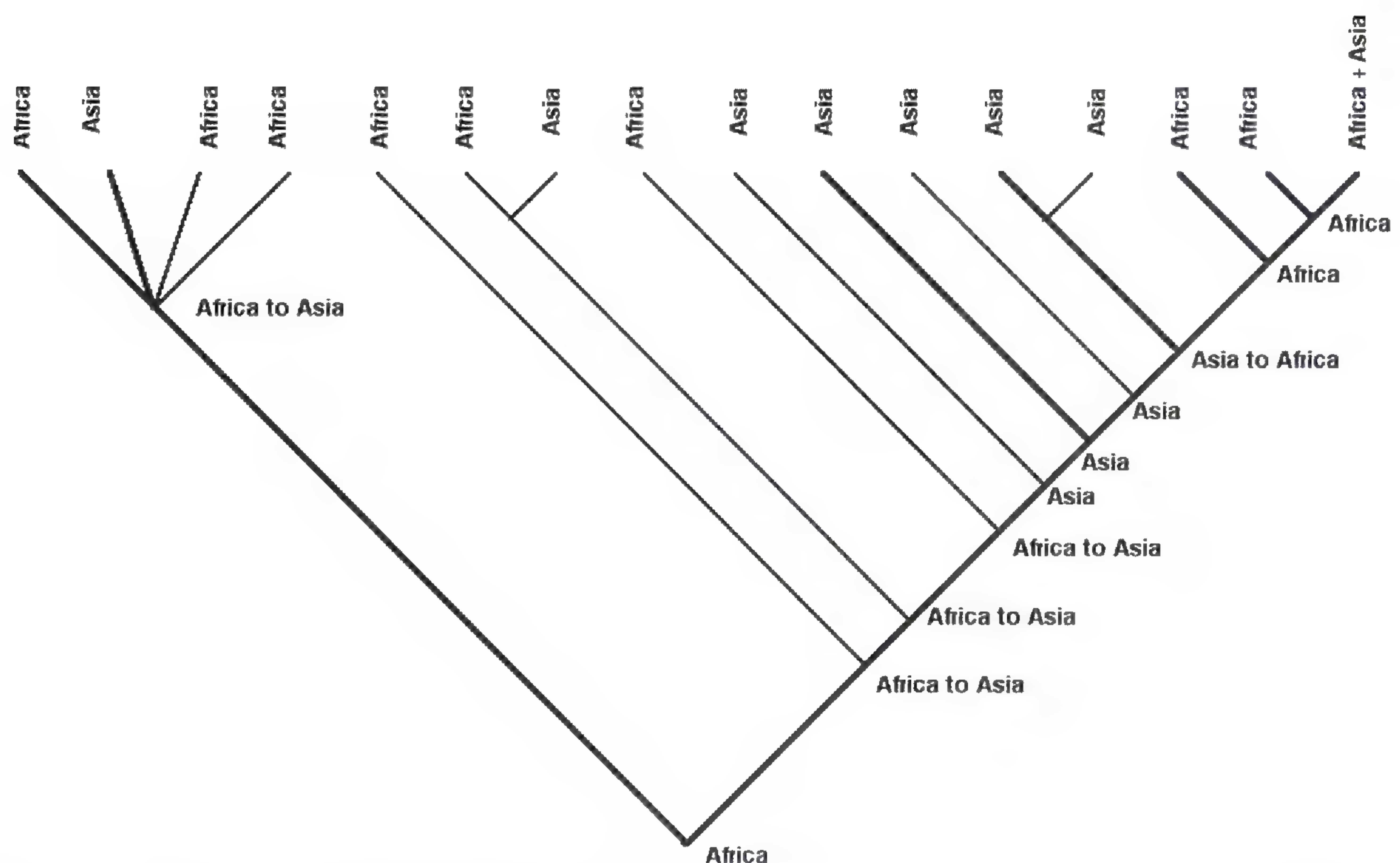


Figure 13. Area cladogram produced by PACT, indicating geographic context of speciation events implied by phylogenies for *Enterobius* and *Oeoshagostomum* (*Conoweberia*). Heavy lines indicate area associations shown by both parasite clades; thin lines indicate area associations shown by only one clade. Areas at nodes indicate episodes of isolation (in either “Asia” or “Africa”) or of biotic expansion (“Africa to Asia,” “Asia to Africa,” or “Asia and Africa”).

the parasite becomes an EID of the newly acquired host.

Brooks (1985) first called attention to analytical similarities in phylogenetic studies of coevolution and biogeography. Given this, it is fairly straightforward to integrate information from both types of studies to obtain some understanding of the evolutionary basis of EIDs. When the background is one of alternating episodes of vicariance and biotic expansion, therefore, episodes of rapid host switching associated with the biotic expansion events may result. The biogeographic analysis of *Enterobius* and *Oesophagostomum* (*Conoweberia*) (Fig. 13) suggests a history of speciation involving alternating episodes of isolation in, and then movements between, Africa and Asia. Furthermore, five of the eight (63%) host switches discovered in Figure 12 occurred during periods of biotic expansion discovered in Figure 13. The exceptions are the switch from primates to rodents in the common ancestor of the parasites *O. xeri* Ortlepp (1922) and *O. susannae* Leroux (1940), which took place in Africa, as well as *O. blanchardi* Travassos & Vogelsang (1932) in orangutans and *O. railletii* Travassos & Vogelsang (1932) in gibbons (perhaps representing a case of sympatric speciation), which took place in Asia. Together, these observations suggest that these host switches, all of which would have been called EIDs at the time of their origins, with the possible exception of the switch from primates to rodent by the common ancestor of *O. xeri* and *O. susannae*, resulted from changes in the

apparency arena rather than the evolution of novel capabilities for host-utilization (Brooks & Hoberg, 2006, 2007a, b).

EIDs may also arise from host switches mediated by ecological fitting within the area of origin when climate changes alter trophic structure. For example, when ancestral humans moved out of the African forest and onto the savannah during the late Pliocene and early Pleistocene, they made a rapid transition from herbivory to facultative carnivory to active predation. During that time, humans apparently shared more than just food with other apex carnivores, becoming hosts to species of cestodes in the genus *Taenia* L. (1758), whose closest relatives inhabit hyenas, large cats, and African hunting dogs (Hoberg et al., 2000, 2001). Thus, colonization of hominids was driven by shared trophic structure in foraging guilds where apex carnivores and ancestral humans foraged on an array of antelope prey that served as intermediate hosts for species of *Taenia*.

Until now, PACT has been used to assess biogeographic and coevolutionary associations among species representing multiple clades. Here, for the first time, we demonstrate how PACT can be used to help further studies of species, speciation, and phylogenetics. Because these are new proposals for applications of Assumption 0 analysis, we begin with the discussion of basic principles.

According to Darwin, the origin of species is historically contingent in space and time (Darwin, 1872). Darwinian thought has always considered

species to be complex systems, whether we consider them to be communities of descent (Darwin, 1872), historical lineages (Simpson, 1944; Wiley, 1978, 1981), cladogenetic units (Hennig, 1966), or historically cohesive information systems irreversibly split from other such units (Brooks & Wiley, 1988; Kornet, 1993; Maynard Smith & Szathmary, 1995) through a variety of cohesion-breaking mechanisms (Brooks & McLennan, 2002). Evolutionary phenomena affecting and producing species operate at the level of reproductive and ecological interactions among individual members of various populations that comprise them.

TOKOGENETICS AND PHYLOGEOGRAPHY

One of the most fundamental assertions made by Hennig (1966) was the distinction between “tokogeny” and “phylogeny.” He asserted that tokogenetic, or within-species, relationships could not be analyzed using phylogenetic systematic principles due to reticulated breeding relationships among organisms, but that phylogenetic, or among-species, relationships could be analyzed using his method because there were no reticulated relationships in phylogeny. By implication, clades that have experienced episodes of hybrid speciation are excluded from Hennigian analysis. The majority of biologists have accepted this self-imposed limitation until recently (see, e.g., Posada & Crandall, 2001).

A “new” guiding principle. In discussing empirical criteria for demarcation among species, Sites and Marshall (2003, 2004) followed Hennig (1966) and Kornet (1993) in suggesting that populations showing reticulated relationships should be considered a single species, whereas populations not showing reticulated relationships should be considered more than one species. Similarly, Riddle and Hafner (2004) suggested that area reticulations should be more common when dealing with populations of the same species than when dealing with populations of different species. The developments of Assumption 0 and of PACT demonstrate that phylogenetic systematic reasoning can be applied to complex historical systems whose members have reticulated relationships. Therefore, viewed in the light of complexity studies, phylogeography (Avice, 2000) is the study of the history of tokogenetic relationships.

A new research program for tokogenetic analyses. Within the realm of an ontologically complex research program, phylogeographic analyses of entities (populations) that show tokogenetic relationships use the following two steps:

1. PACT analysis of all available data. Convert all gene trees into taxon-area cladograms (Fig. 14). This is accomplished by replacing the names of the haplotypes distinguished by each gene tree with the areas they inhabit. Ideally, both mitochondrial and nuclear genes should be used, the former because they retain patterns of introgression well, and the latter because they provide direct evidence of gene flow. Then combine the taxon-area cladograms using PACT. This produces a GAC (Fig. 15) depicting the geographic context of the deployment of genetic information within and among the geographically defined populations.

2. Assessing species boundaries and patterns of genetic cohesion. As suggested by Sites and Marshall (2003, 2004), the Hennigian distinction between tokogeny and phylogeny provides a straightforward empirical criterion for determining how many different species are represented in a sample of populations. Species (macro-species of Kornet, 1993) are identified by the largest inclusive subset (clade) of populations (micro-species of Kornet, 1993) that show reticulated relationships. Using this formalism, Figure 15 depicts three species, comprising populations a–d, populations e–h, and populations i–l, respectively.

The next step is to ascertain the patterns of genetic cohesion among the populations within each species. A truly panmictic species would be characterized by all populations showing reticulated relationships with each other, producing a GAC in the form of a single unresolved polytomy for all populations. In a Darwinian world, historical contingencies of heredity (the nature of the organism) and circumstances (the nature of the conditions) should produce various degrees of within-species cohesion. The hypothetical three-species system in Figure 15 illustrates this point. Each species is characterized by four populations. Species 1 (populations a–d) has one unique population (d) and three populations exhibiting a total of four reticulations (one each for populations a and c, two for population b), combining for eight inter-population connections. Species 2 (populations e–h) has two unique populations (f and h) and two populations exhibiting a total of two duplications (one each for e and g), combining for 13 inter-population connections. Finally, species 3 (populations i–l) has one unique population (j) and three populations exhibiting a total of four duplications (one each for k and l, two for i), combining for 14 inter-population connections. If the magnitude of gene flow is comparable in all inter-population connections, then species 2 and 3 would be deemed more cohesive than species 1, and species 2 might be considered to have achieved its high degree of cohesion more

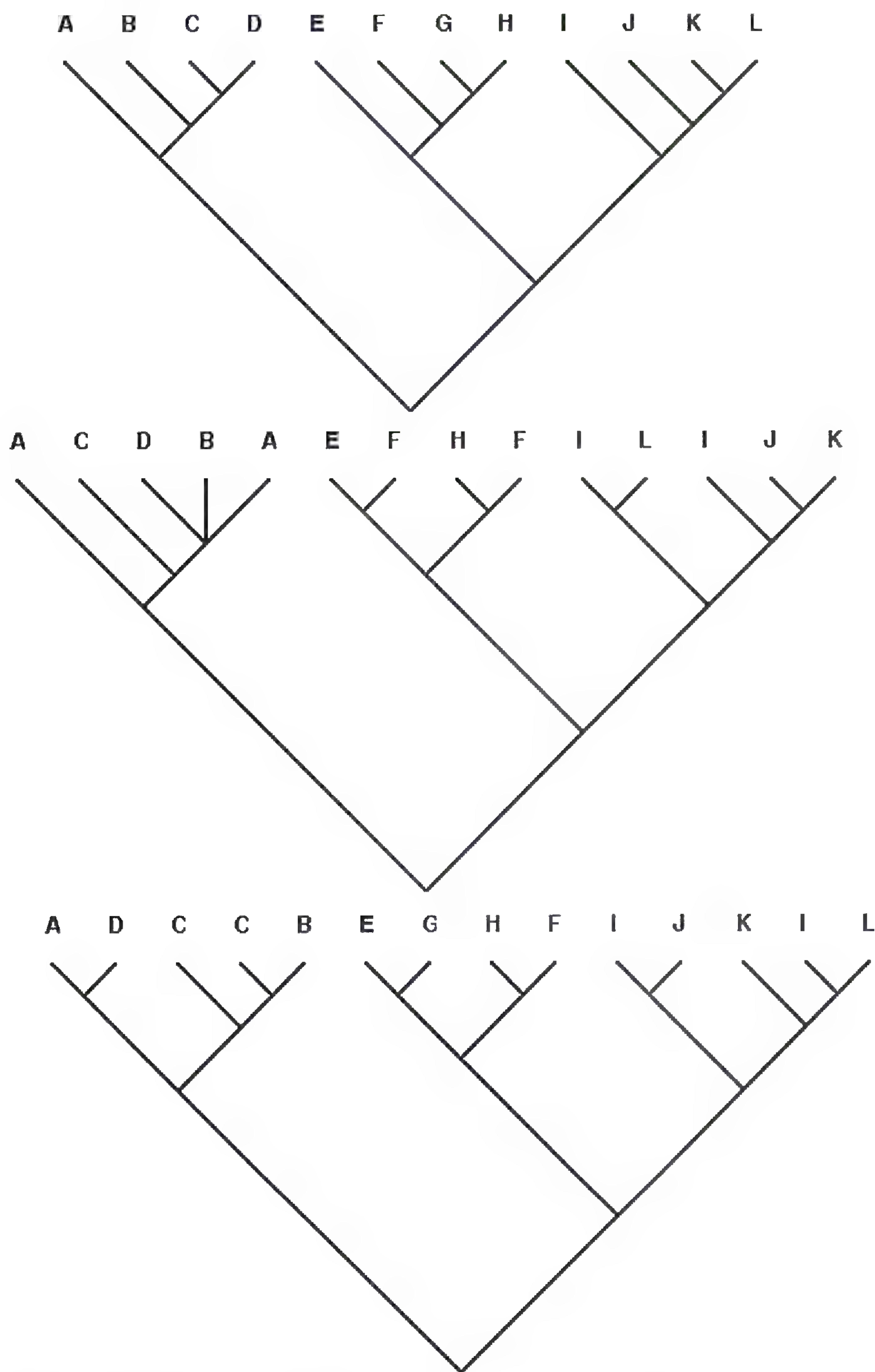


Figure 14. PACT and phylogeography 1. Three gene trees for various populations.

efficiently than species 3, needing only two duplications to achieve 13 inter-population connections as opposed to four duplications to achieve 14 inter-population connections.

PHYLOGENETICS

The Hennigian assumption that phylogeny can be represented accurately by an internested branching diagram is refuted each time a new species is formed by the hybridization of two (or more) other species. This produces a reticulated phylogenetic history that cannot be represented by a purely internested

branching diagram that is the result of analyses with standard phylogenetic methods.

Most zoologists consider hybrid speciation to be rare in animals, but botanists have recognized for a long time that many plant species have been formed in this manner (Vriesendorp & Bakker, 2005 and examples therein). Furthermore, the endosymbiont theory of the origin of eukaryotes (e.g., Schwartz & Dayhoff, 1978; Margulis, 1981; Sitte, 1993) postulates that certain major transitions in evolution (sensu Maynard Smith & Szathmary, 1995) are the result of hybridization. At least some prokaryotes have an interesting ability that refers to the capability of

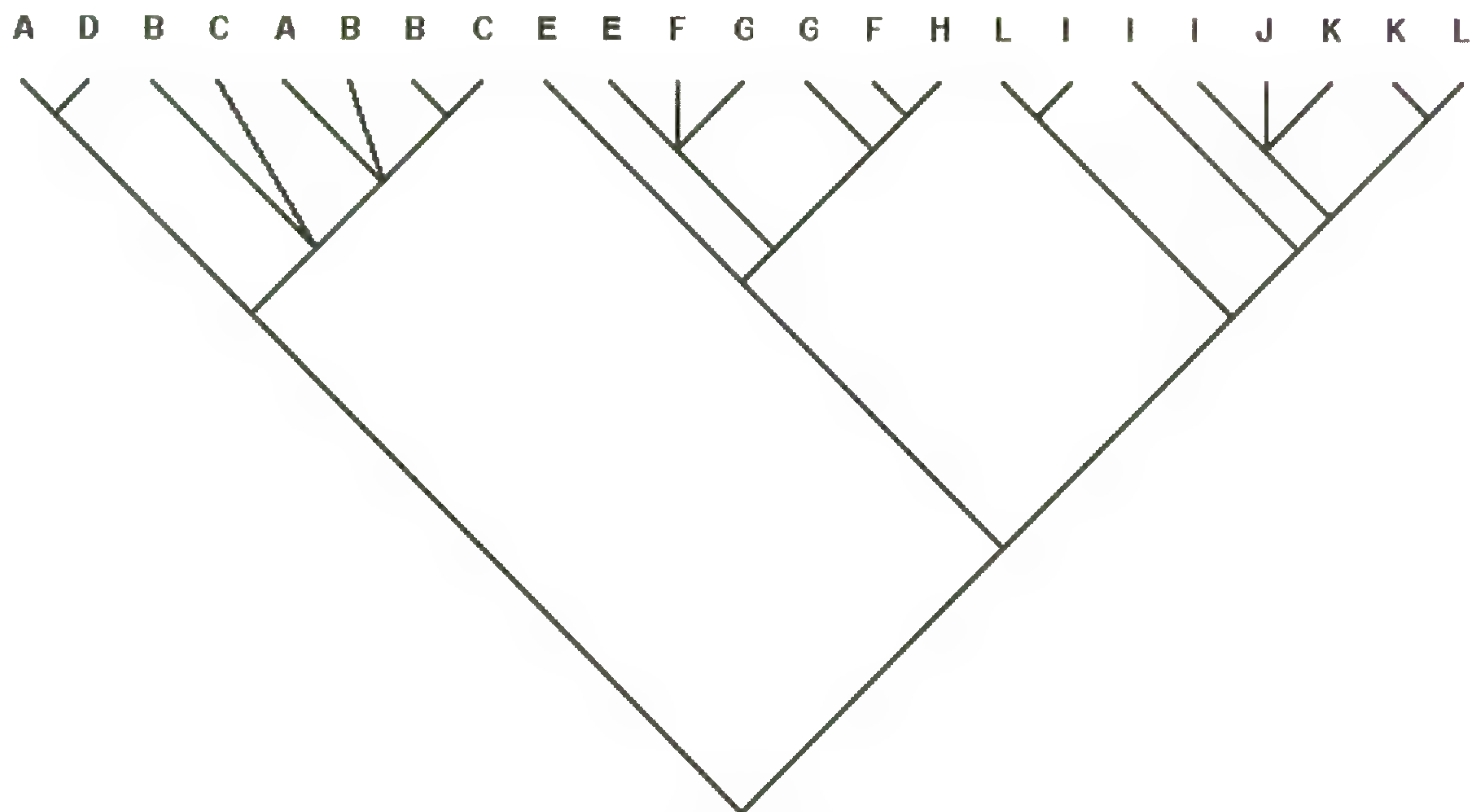


Figure 15. PACT and phylogeography 2. PACT analysis of three gene trees shown in Figure 14.

producing reticulated phylogenetic relationships. In times of stress, these prokaryotes not only shut down their metabolic activities but may actually jettison the portion of their genomes associated with metabolic activities. Furthermore, they are capable of regaining genetic machinery for metabolic activities, but not always from a close relative.

Partial hybridization is a common process for prokaryotes (Woese, 2000, 2002). In bacteria, genes for resistance to antibiotics are usually situated on plasmids. These plasmids behave as accessory bacterial chromosomes and can be transferred by bacterial conjugation. Their plasmids can be exchanged via a tubular connection between two bacterial cells, and resistance to antibiotics can become widespread to relative as well as nonrelative species. Another way of transferring genetic material between two species of bacteria can be accomplished by transduction by a bacterial virus (bacteriophages or phage). This partial hybridization by exchange of genetic material, however, is not only restricted to bacteria. In case of bacterial transformation (by, e.g., *Agrobacterium tumefaciens* [Smith & Townsend, 1907] Conn [1942] or *A. rhizogenes* [Riker et al., 1930] Conn [1942]), genetic material may even be transferred from a prokaryote to a higher plant species.

A “new” guiding principle. Phylogeneticists were aware early on that species of hybrid origin might pose special problems (Bremer & Wanntorp, 1979; Funk, 1981), and this was the subject of a symposium at the second meeting of the Willi Hennig Society (Humphries, 1983; Nelson, 1983; Wagner, 1983; Wanntorp, 1983). Funk (1985), however, was the first to propose that analyzing correlated homoplasies in analyses of clades containing species of hybrid origin

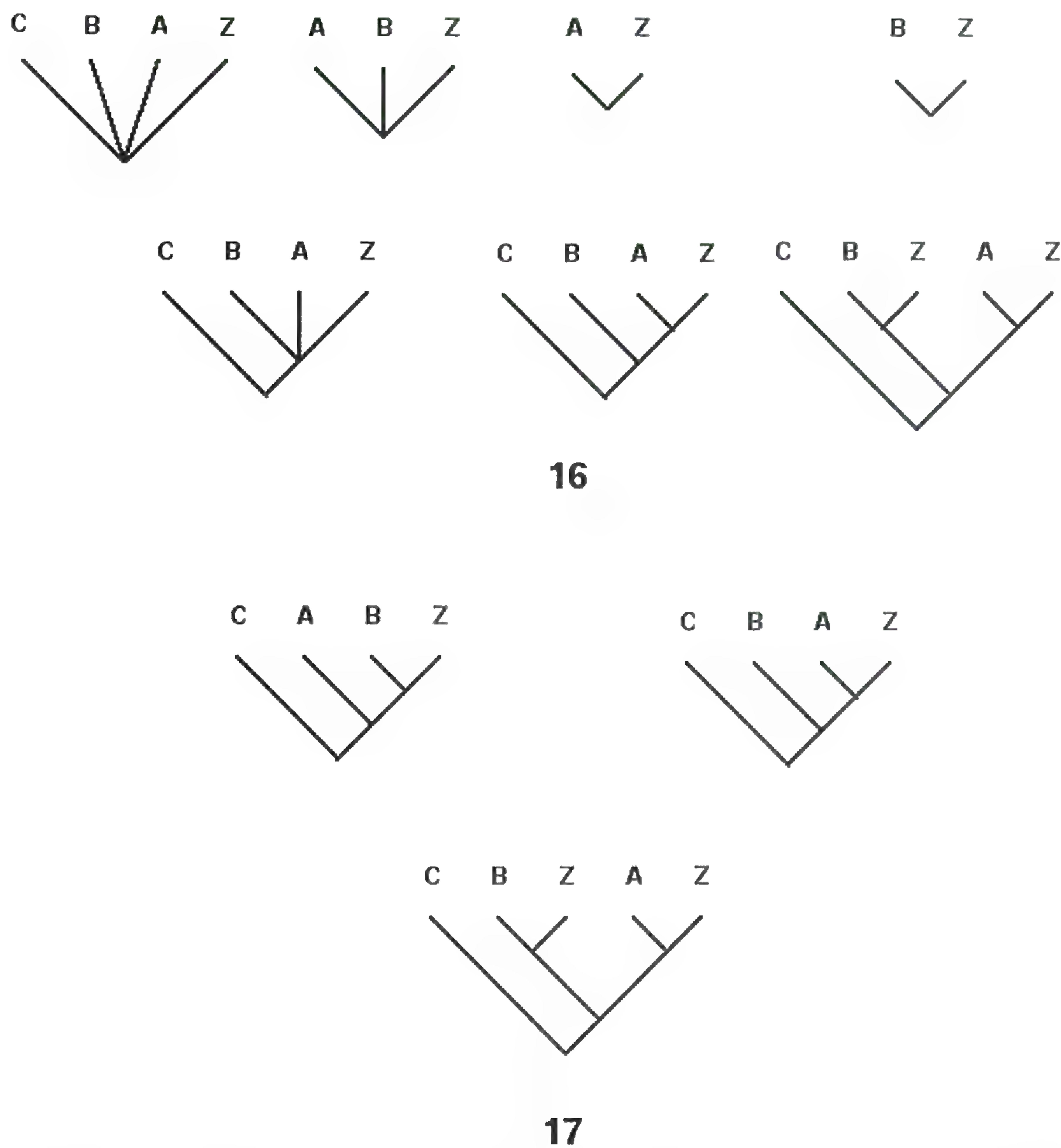
could be used to detect the hybrid species. She noted that when there are multiple equally parsimonious trees for a clade suspected of containing species of hybrid origins, those trees often represent a small number of distinct topologies. Furthermore, inspection of the trees indicated that most of the homoplasy and, thus, most of the ambiguity was contributed by a small number of species. Removal of those species from the analysis produced a stable topology and reduced homoplasy. Funk’s insights, combined with PACT, form the basis for an approach to phylogenetic analysis under Assumption 0.

A new research program for phylogenetic analyses. Within the realm of an ontologically complex research program, phylogenetic analyses of entities show cladogenetic as well as reticulated relationships, and the following two steps are applied.

Step 1—PACT analysis of all available data. Convert all transformation series into character-state trees, then combine them using PACT (Figs. 16, 17), ignoring the plesiomorphic state. This produces a phylogenetic tree in which all traits interpreted as homoplasy by standard Hennigian analysis will be interpreted as homologies resulting from horizontal transfer/hybridization.

Step 2—Explaining reticulations. There are two different, yet intertwined, issues involved in this step. The first is differentiating reticulated homology from true homoplasy. The second is differentiating horizontal transfer from hybrid speciation.

The key to deciding the first issue is having phylogenies based on as many characters and character types as possible—comprehensive total-evidence studies. We would then expect, as suggested by Funk (1985), that cases of hybrid speciation would



Figures 16 and 17. —16. PACT and analysis of hybrid species. Top four are binary character-state trees indicating distribution of apomorphic states. Bottom three are PACT results of (from left) combining 1 and 2, then (1 + 2) and 3, then (1 + 2 + 3) and 4. —17. PACT and analysis of hybrid species. Top 2 diagrams are gene trees; bottom diagram is PACT result for combining the gene trees. Parental species are sister species.

appear in a PACT analysis as duplications supported by many characters (because hybrid speciation involves combinations of whole genomes), whereas cases of true homoplasy would appear as duplications supported by single characters (because homoplasy involves single characters). For such cases, relying on epistemological parsimony would have an ontological justification. Consider this thought experiment:

One species (Z) forms a sister relationship with each of two other species (A and B) in a PACT analysis. If you have one case of (BZ) and one of (AZ), there are three possible explanations: (1) hybrid speciation (Assumption 0 default); (2) one case of horizontal transfer (A is the true sister species of Z, with horizontal transfer to B or vice versa); or (3) one case of true homoplasy (independent origin). Using epistemological parsimony alone, there is no way to choose, based on the evidence at hand.

Next, suppose you have five cases of (BZ) and five cases of (AZ). Now the explanations become (1) one case of hybrid speciation; (2) five cases of horizontal transfer or one to four cases of horizontal transfer of linked traits; and (3) five cases of independent evolution or one to four cases of “co-adapted trait complex” or “adaptive syndrome” evolution. In this case, epistemological parsimony would lead us to choose hybrid speciation unambiguously as the preferred explanation. That means we hypothesize that all shared similarities among the species are homologous.

This approach is not guaranteed to produce correct results, of course. It works when enough of the heritage of each parental species is expressed in the hybrid product(s). We expect that the more distantly related the parents are, the less likely we will be able to recover the parental species, but also the less likely

that hybrid speciation will occur. In addition, we often expect hybrid species to segregate phylogenetically more with one parent than another (see McDade, 1990, 1992; Mindell, 1992, 1993; Zrzavy & Skala, 1993; Skala & Zrzavy, 1994). For cases in which ambiguity in phylogenetic analysis results from a small number of species and when there is abundant evidence of correlated homoplasy, Funk's approach gives us an excellent first approximation as to the possibility that the ambiguity is due to those species being hybrids, justifying the use of Assumption 0 analysis. However, when larger numbers of species are involved, the explanations for character ambiguity may become very complex.

For example, recent studies of prokaryote phylogenetics have raised the possibility that this propensity on the part of prokaryotes has produced some wildly reticulated phylogenetic relationships. Some have even expressed the opinion that this means we should change our view of phylogeny, from a tree to a reticulated network (Doolittle, 2000). Maynard Smith and Szathmary (1995), however, pointed out that genetic machinery associated with storing and transmitting information shows no similar evidence of horizontal transfer. Thus, the true phylogeny may still be tree-like in pattern, although substantial homoplasy in metabolic genes may be present due to horizontal transfer.

Regardless of the eventual outcome of this debate, we do not have to develop fundamentally different approaches to assessing the evidence about relationships. It simply means that the more horizontal transfer there has been, the more we will have to resort to species duplications to find the source of those transfers. Much more work, including substantial modeling effort, is needed to help determine when it is appropriate to use Assumption 0 in phylogenetic analysis and when it is not.

We will now address the second issue. When Assumption 0 and PACT are applied in this context, all reticulations are taken as evidence of introgression, but introgression can produce multiple outcomes, of which speciation is only one. Differentiating the various forms and outcomes of introgression requires a robust and unified ontological species concept (the evolutionary species concept; Simpson, 1944, 1951; Wiley, 1978) as well as a broad-based epistemological view of the diversity of kinds of species. As we suggested earlier, the PACT analysis is the beginning of an explanation, not an explanation in itself. Inference of species boundaries requires additional tests (such as nested clade analysis) to highlight the independent lineages and experiments to determine the strength of cohesion of those lineages (for a discussion and protocol, see Brooks & McLennan, 2002).

Finally, we come to the issue of how to incorporate reticulated phylogenetic events into our classifications. There are two choices. Do we classify a tree with reticulations, or do we try to classify a reticulated network? If we choose the former, hybrid species will appear in two different clades, reflecting their phylogenetic origins. This disturbs traditional classifications in which each taxon has a unique place. If we choose to try to classify a network, each species of hybrid origin would be classified separately from either parental species, so the resulting classification would maintain the tradition of each taxon having a unique placement, but could not provide information about the phylogenetic relationships of the hybrid species. If we wish our classifications to reflect what we think we know about evolution, it seems that we will have to opt for the first alternative.

CONCLUSIONS: BACK TO THE FUTURE

Evolution is historically contingent and complex. Darwinism provides an ontology of complexity. Historical contingency requires phylogenetic information to document and explain patterns. Representing this complexity so that we have some understanding of what needs to be explained requires a method that allows reticulated relationships to be found. A new algorithm, PACT (Wojcicki & Brooks, 2004, 2005), has been developed for that purpose. Applying this algorithm in historical biogeographical, coevolutionary, phylogeographical, and phylogenetic studies invokes Assumption 0; all data must be analyzed without modification. Thereby, the resulting patterns of related areas, species, and populations contain duplications of entities to such an extent that these patterns are logically consistent with all input data. Further inference of these duplications asks for an ontology of complexity in which multiple evolutionary processes resulting in both divergent and reticulated patterns nest comfortably.

Methods for phylogenetic analysis that only deal with divergent patterns and do not allow duplication of entities disallow complex evolutionary processes for explaining the convergent (reticulated) patterns and thereby oversimplify the historical relationships among the entities of analysis. Furthermore, several studies mentioned in this paper show that several older and newer concepts like the taxon pulse (Erwin, 1979), ecological fitting (Janzen, 1985b), phylogenetic analysis of hybrids (Funk, 1985), and the oscillation hypothesis (Janz & Nylin, 2007) allow for explanation of reticulated patterns.

In order to perform comparative analyses based upon an ontology of complexity, we do not have to begin again because Darwinism is based on such an ontology. We do, however, have to move forward,

beyond the oversimplification of evolution associated with much of 20th-century evolutionary biology. Basic principles of phylogenetic systematics, coupled with Assumption 0, can accommodate analyses of complex evolutionary associations at various levels of organization. Once we begin doing this, we will recapture earlier ideas and integrate them into a renewed appreciation for the Darwinian panorama.

Literature Cited

- Adamson, M. L. & J. N. Caira. 1994. Evolutionary factors influencing the nature of parasite specificity. *Parasitol.* (London) 109: S85–S95.
- Avice, J. C. 2000. *Phylogeography: The History and Formation of Species*. Harvard Univ. Press, Cambridge.
- Bouchard, P., D. R. Brooks & D. K. Yeates. 2004. Mosaic macroevolution in Australian wet tropics arthropods: Community assemblage by taxon pulses. Pp. 425–469 in C. Moritz & E. Bermingham (editors), *Rainforests: Past, Present, Future*. Univ. of Chicago Press, Chicago.
- Bowler, P. J. 1983. *The Eclipse of Darwinism*. Johns Hopkins Univ. Press, Baltimore.
- Bremer, K. & H.-E. Wanntorp. 1979. Hierarchy and reticulation in systematics. *Syst. Zool.* 28: 624–627.
- Brooks, D. R. 1985. Historical ecology: A new approach to studying the evolution of ecological associations. *Ann. Missouri. Bot. Gard.* 72: 660–680.
- . 1998. The unified theory of evolution and selection processes. Pp. 113–128 in G. van de Vijver, S. N. Salthe & M. Delpo (editors), *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-Organization*. Kluwer Academic Publishers, Dordrecht.
- . 2000. The nature of the organism: Life takes on a life of its own. *Proc. New York Acad. Sci.* 901: 257–265.
- . 2001. Evolution in the information age: Rediscovering the nature of the organism. *Semiotics Evol. Energy Developm.* 1. Available at: <<http://www.library.utoronto.ca/see>>, accessed 17 March 2008.
- . 2002. Taking evolutionary transitions seriously. *Semiotics Evol. Energy Developm.* 2: 6–24, Available at: <<http://www.library.utoronto.ca/see>>, accessed 17 March 2008.
- & E. O. Wiley. 1986. *Evolution as Entropy: Toward a Unified Theory of Biology*, 1st ed. Univ. of Chicago Press, Chicago.
- & ———. 1988. *Evolution as Entropy: Toward a Unified Theory of Biology*, 2nd ed. Univ. of Chicago Press, Chicago.
- & D. A. McLennan. 1991. *Phylogeny, Ecology and Behavior: A Research Program in Comparative Biology*. Univ. of Chicago Press, Chicago.
- & ———. 1993. *Parascript: Parasites and the Language of Evolution*. Smithsonian Institution Press, Washington, D.C.
- & ———. 2000. The nature of the organism and the emergence of selection processes and biological signals. Pp. 185–218 in E. Taborsky (editor), *Semiotics, Evolution, Energy*. Shaker Verlag, Aachen, Germany.
- & ———. 2001. A comparison of a discovery-based and an event-based method of historical biogeography. *J. Biogeogr.* 28: 757–767.
- & ———. 2002. *The Nature of Diversity: An Evolutionary Voyage of Discovery*. Univ. of Chicago Press, Chicago.
- & ———. 2003. Extending phylogenetic studies of coevolution: Secondary Brooks parsimony analysis, parasites, and the Great Apes. *Cladistics* 19: 104–119.
- & A. L. Ferrao. 2005. The historical biogeography of coevolution: Emerging infectious diseases are evolutionary accidents waiting to happen. *J. Biogeogr.* 32: 1291–1299.
- & K. E. Folinsbee. 2005. Paleobiogeography: Documenting the ebb and flow of evolutionary diversification. Pp. 15–43 in B. S. Lieberman & A. Stigall Rode (editors), *Paleontological Society Papers*, Vol. 11.
- & E. P. Hoberg. 2006. Systematics and emerging infectious diseases: From management to solution. *J. Parasitol.* 92: 426–429.
- & ———. 2007a. How will global climate change affect parasite-host assemblages? *Trends Parasitol.* 23: 571–574.
- & ———. 2007b. Darwin's necessary misfit and the sloshing bucket: The evolutionary biology of emerging infectious diseases. *Evol. Outreach Educ.* 1: 2–9.
- Brues, C. T. 1920. The selection of food-plants by insects, with special reference to lepidopterous larvae. *Amer. Midl. Naturalist* 54: 313–332.
- . 1924. The specificity of food-plants in the evolution of phytophagous insects. *Amer. Midl. Naturalist* 58: 127–144.
- Collier, J. & C. Hooker. 1999. Complexly organised dynamical systems. *Open Syst. Info. Dynamics* 6: 241–302.
- Cracraft, J. 1985. Biological diversification and its causes. *Ann. Missouri. Bot. Gard.* 72: 794–822.
- Darlington, P. J. 1943. Carabidae of mountains and islands: Data on the evolution of isolated faunas, and on atrophy of wings. *Ecol. Monogr.* 13: 37–61.
- Darwin, C. 1872. *The Origin of Species*, 6th ed. John Murray, London.
- Doolittle, W. F. 2000. Uprooting the tree of life. *Sci. Amer.* 282: 90–95.
- Dowling, A. P. G. 2002. Testing the accuracy of TreeMap and Brooks parsimony analyses of coevolutionary patterns using artificial associations. *Cladistics* 18: 416–435.
- , M. G. P. van Veller, E. P. Hoberg & D. R. Brooks. 2003. A priori and a posteriori methods in comparative evolutionary studies of host-parasite associations. *Cladistics* 19: 240–253.
- Ehrlich, P. R. & P. H. Raven. 1964. Butterflies and plants: A study in coevolution. *Evolution* 18: 586–608.
- Eldredge, N. 1985. *Unfinished Synthesis: Biological Hierarchies and Modern Evolutionary Thought*. Oxford Univ. Press, New York.
- . 1986. Information, economics and evolution. *Ann. Rev. Ecol. Syst.* 17: 351–369.
- . 1989. *Macroevolutionary Dynamics: Species, Niches and Adaptive Peaks*. McGraw-Hill, New York.
- Endler, J. A. 1982. Problems in distinguishing historical from ecological factors in biogeography. *Amer. Zool.* 22: 441–452.
- Erwin, T. L. 1979. Thoughts on the evolutionary history of ground beetles: Hypotheses generated from comparative faunal analyses of lowland forest sites in temperate and tropical regions. Pp. 539–592 in T. L. Erwin, G. E. Ball & D. R. Whitehead (editors), *Carabid Beetles—Their Evolution, Natural History, and Classification*. W. Junk, The Hague.
- . 1981. Taxon pulses, vicariance, and dispersal: An evolutionary synthesis illustrated by carabid beetles. Pp. 159–196 in G. Nelson & D. E. Rosen (editors), *Vicariance Biogeography—A Critique*. Columbia Univ. Press, New York.

- Fahrenholz, H. 1913. Ectoparasiten unde abstammungslehre. *Zool. Anz. (Leipzig)* 41: 371–374.
- Folinsbee, K. & D. R. Brooks. 2007. Early hominoid biogeography: Pulses of dispersal and differentiation. *J. Biogeogr.* 34: 383–397.
- Funk, V. A. 1981. Special concerns in estimating plant phylogenies. Pp. 73–86 in V. A. Funk & D. R. Brooks (editors), *Advances in Cladistics*, Vol. I. New York Botanical Garden, New York.
- . 1985. Phylogenetic patterns and hybridization. *Ann. Missouri Bot. Gard.* 72: 681–715.
- Gould, S. J. 1983. The hardening of the modern synthesis. Pp. 75–76 in M. Grene (editor), *Dimensions of Darwinism: Themes and Counterthemes in Twentieth-Century Evolutionary Theory*. Cambridge Univ. Press, Cambridge.
- Green, M. D., M. G. P. van Veller & D. R. Brooks. 2002. Assessing modes of speciation: Range asymmetry and biogeographical congruence. *Cladistics* 18: 112–124.
- Halas, D., D. Zamparo & D. R. Brooks. 2005. A protocol for studying biotic diversification by taxon pulses. *J. Biogeogr.* 32: 249–260.
- Heaney, L. R. 2000. Dynamic disequilibrium: A long-term, large-scale perspective on the equilibrium model of island biogeography. *Global Ecol. Biogeogr.* 9: 59–74.
- Hennig, W. 1966. *Phylogenetic Systematics*. Univ. of Illinois Press, Urbana.
- Hoberg, E. P. & A. Adams. 2000. Phylogeny, history and biodiversity: Understanding faunal structure and biogeography in the marine realm. *Bull. Scand. Soc. Parasitol.* 10: 19–37.
- , A. Jones, R. L. Rausch, K. S. Eom & S. L. Gardner. 2000. A phylogenetic hypothesis for species of the genus *Taenia* (Eucestoda: Taeniidae). *J. Parasitol.* 86: 89–98.
- , N. Alkire, A. de Queiroz & A. Jones. 2001. Out of Africa: Origins of the *Taenia* tapeworms in humans. *Proc. Roy. Soc. London, Ser. B, Biol. Sci.* 268: 781–787.
- Humphries, C. J. 1983. Primary data in hybrid analysis. Pp. 89–103 in N. I. Platnick & V. A. Funk (editors), *Advances in Cladistics II*. Columbia Univ. Press, New York.
- & L. R. Parenti. 1999. *Cladistic Biogeography: Interpreting Patterns of Plant and Animal Distributions*, 2nd ed. Oxford Univ. Press, New York.
- Janzen, N. & S. Nylin. 2007. The oscillation hypothesis of host plant-range and speciation. Pp. 203–215 in K. J. Tilmon (editor), *Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects*. Univ. of California Press, Berkeley.
- , ——— & N. Wahlberg. 2006. Diversity begets diversity: Host expansions and the diversification of plant-feeding insects. *BMC Evol. Biol.* 6: 4.
- Janzen, D. H. 1968. Host plants as islands in evolutionary and contemporary time. *Amer. Midl. Naturalist* 102: 592–595.
- . 1973a. Host plants as islands: Competition in evolutionary and contemporary time. *Amer. Midl. Naturalist* 107: 786–790.
- . 1973b. Comments on host-specificity of tropical herbivores and its relevance to species richness. Pp. 201–211 in V. H. Heywood (editor), *Taxonomy and Ecology*. Academic Press, New York.
- . 1980. When is it coevolution? *Evolution* 34: 611–612.
- . 1981. Patterns of herbivory in a tropical deciduous forest. *Biotropics* 13: 271–282.
- . 1983. Dispersal of seeds by vertebrate guts. Pp. 232–262 in D. J. Futuyma & M. Slatkin (editors), *Coevolution*. Sinauer Associates, Sunderland, Massachusetts.
- . 1985a. Coevolution as a process: What parasites of animals and plants do NOT have in common. Pp. 83–99 in K. C. Kim (editor), *Coevolution of Parasitic Arthropods and Mammals*. Wiley and Sons, New York.
- . 1985b. On ecological fitting. *Oikos* 45: 308–310.
- & P. S. Martin. 1982. Neotropical anachronisms: The fruits the Gomphotheres ate. *Science* 215: 19–27.
- Kellogg, V. L. 1896. New Mallophaga, I. With special reference to a collection from maritime birds of the Bay of Monterey, California. *Proc. Calif. Acad. Sci.* 6: 31–168.
- . 1913. Distribution and species-forming of ectoparasites. *Amer. Midl. Naturalist* 47: 129–158.
- Kornet, D. J. 1993b. Permanent splits as speciation events: A formal reconstruction of the internodal species concept. *J. Theor. Biol.* 164: 407–435.
- Lieberman, B. S. 2000. *Paleobiogeography*. Plenum/Kluwer Academic, New York.
- . 2003a. Unifying theory and methodology in biogeography. *Evol. Biol.* 33: 1–25.
- . 2003b. Paleobiogeography: The relevance of fossils to biogeography. *Ann. Rev. Ecol. Syst.* 34: 51–69.
- Losos, J. B. & D. Schluter. 2000. Analysis of an evolutionary species-area relationship. *Nature* 408: 847–850.
- MacArthur, R. H. & E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. *Evolution* 17: 373–387.
- & ———. 1967. *The Theory of Island Biogeography*. Princeton Univ. Press, Princeton, New Jersey.
- Margulis, L. 1981. *Symbiosis in Cell Evolution*. W. H. Freeman, San Francisco.
- Marshall, C. J. & J. K. Liebherr. 2000. Cladistic biogeography of the Mexican transition zone. *J. Biogeogr.* 27: 203–216.
- Maynard Smith, J. & E. Szathmary. 1995. *The Major Transitions in Evolution*. W. H. Freeman, San Francisco.
- McDade, L. A. 1990. Hybrids and phylogenetic systematics I. Patterns of character expression in hybrids and their implications for cladistic analysis. *Evolution* 44: 1685–1700.
- . 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346.
- McLennan, D. A. & D. R. Brooks. 2002. Complex histories of speciation and dispersal: An example using some Australian birds. *J. Biogeogr.* 29: 1055–1066.
- Mindell, D. P. 1992. Phylogenetic consequences of symbioses: Eukarya and Eubacteria are not monophyletic taxa. *BioSystems* 27: 53–62.
- . 1993. Merger of taxa and definition of monophyly. *BioSystems* 31: 130–133.
- Mode, C. J. 1958. A mathematical model for the co-evolution of obligate parasites and their hosts. *Evolution* 12: 158–165.
- Nelson, G. 1983. Reticulation in cladograms. Pp. 105–111 in N. I. Platnick & V. A. Funk (editors), *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society*. Columbia Univ. Press, New York.
- Nylin, S. & N. Janz. 2007. Butterfly host plant range: An example of plasticity as a promoter of speciation? *Evol. Ecol.* DOI: 10.1007/s10682-007-9205-5.
- Page, R. D. M. (editor). 2002. *Tangled Trees*. Univ. of Chicago Press, Chicago.
- Posada, D. & K. A. Crandall. 2001. Intraspecific gene genealogies: Trees grafting into networks. *Trends Ecol. Evol.* 16: 37–45.
- Riddle, B. R. & D. J. Hafner. 2004. The past and future roles of phylogeography in historical biogeography. Pp. 93–110 in M. V. Lomolino & L. R. Heaney (editors), *Frontiers in Biogeography: New Directions in the Geography of Nature*. Sinauer Associates, Sunderland, Massachusetts.

- Salvador Arias, J., I. J. Garzon-Orduna, F. Lopez-Osorio, E. Parada-Vargas & D. Rafael Miranda-Esquivel. 2008. What is PACT really? *Cladistics* 24: 1–12.
- Schwartz, R. M. & M. O. Dayhoff. 1978. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* 199: 395–403.
- Simberloff, D. S. 1987. Calculating probabilities that cladograms match: A method of biogeographical inference. *Syst. Zool.* 36: 56–59.
- , K. L. Heck, E. D. McCoy & E. F. Connor. 1980. There have been no statistical tests of cladistic biogeographic hypotheses. Pp. 40–63 in G. Nelson & D. E. Rosen (editors), *Vicariance Biogeography: A Critique*. Columbia Univ. Press, New York.
- Simpson, G. G. 1944. *Tempo and Mode in Evolution*. Columbia Univ. Press, New York.
- . 1951. The species concept. *Evolution* 5: 285–298.
- . 1953. *The Major Features of Evolution*. Columbia Univ. Press, New York.
- Sites, J. W., Jr. & J. C. Marshall. 2003. Delimiting species: A renaissance issue in systematic biology. *Trends Ecol. Evol.* 18: 462–470.
- & ———. 2004. Operational criteria for delimiting species. *Ann. Rev. Ecol. Syst.* 35: 199–227.
- Sitte, P. 1993. Symbiogenetic evolution of complex cells and complex plastids. *Eur. J. Protistol.* 29: 131–143.
- Skala, Z. & J. Zrzavy. 1994. Phylogenetic reticulations and cladistics: Discussion of methodological concepts. *Cladistics* 10: 305–313.
- Spironello, M. & D. R. Brooks. 2003. Dispersal and diversification in the evolution of *Inseliellium*, an archipelagic dipteran group. *J. Biogeogr.* 30: 1563–1573.
- Trouvé, S., P. Sasal, J. Jourdans, F. N. Renaud & S. Morand. 1998. The evolution of life-history traits in parasitic and free-living platyhelminths: A new perspective. *Oecologia* 115: 370–378.
- van Veller, M. G. P. & D. R. Brooks. 2001. When simplicity is not parsimonious: A priori and a posteriori methods in historical biogeography. *J. Biogeogr.* 28: 1–11.
- , M. Zandee & D. J. Kornet. 1999. Two requirements for obtaining valid common patterns under different assumptions in vicariance biogeography. *Cladistics* 15: 393–406.
- , D. J. Kornet & M. Zandee. 2000. Methods in vicariance biogeography: Assessment of the implementations of assumptions zero, 1 and 2. *Cladistics* 16: 319–345.
- , M. Zandee & D. J. Kornet. 2001. Measures for obtaining inclusive sets of area cladograms under assumptions zero, 1 and 2 with different methods for vicariance biogeography. *Cladistics* 17: 248–259.
- , D. J. Kornet & M. Zandee. 2002. A posteriori and a priori methodologies for testing hypotheses of causal processes in vicariance biogeography. *Cladistics* 18: 207–217.
- , D. R. Brooks & M. Zandee. 2003. Cladistic and phylogenetic biogeography: The art and the science of discovery. *J. Biogeogr.* 30: 319–329.
- Verschaffelt, E. 1910. The cause determining the selection of food in some herbivorous insects. *Proc. Acad. Sci., Amsterdam* 13: 536–542.
- Vriesendorp, B. & F. T. Bakker. 2005. Reconstructing patterns of reticulate evolution in angiosperms: What can we do? *Taxon* 54: 593–604.
- Wagner, W. H., Jr. 1983. Reticulistics: The recognition of hybrids and their role in cladistics and classification. Pp. 63–79 in N. I. Platnick & V. A. Funk (editors), *Advances in Cladistics II*. Columbia Univ. Press, New York.
- Wanntorp, H. E. 1983. Reticulated cladograms and the identification of hybrid taxa. Pp. 81–88 in N. I. Platnick & V. A. Funk (editors), *Advances in Cladistics II*. Columbia Univ. Press, New York.
- Whittaker, R. J. 2000. Scale, succession and complexity in island biogeography: Are we asking the right questions? *Global Ecol. Biogeogr.* 9: 75–85.
- Wiley, E. O. 1978. The evolutionary species concept reconsidered. *Syst. Zool.* 27: 17–26.
- . 1981. *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. Wiley-Interscience, New York.
- . 1986. Methods in vicariance biogeography. Pp. 283–306 in P. Hovenkamp (editor), *Systematics and Evolution: A Matter of Diversity*. Utrecht Univ. Press, Utrecht.
- . 1988a. Parsimony analysis and vicariance biogeography. *Syst. Zool.* 37: 271–290.
- . 1988b. Vicariance biogeography. *Ann. Rev. Ecol. Syst.* 19: 513–542.
- Wilson, E. O. 1959. Adaptive shift and dispersal in a tropical and fauna. *Evolution* 13: 122–144.
- . 1961. The nature of the taxon cycle in the Melanesian ant fauna. *Amer. Midl. Naturalist* 95: 169–193.
- Woese, C. R. 2000. Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. U.S.A.* 97: 8392–8396.
- . 2002. On the evolution of cells. *Proc. Natl. Acad. Sci. U.S.A.* 99: 8742–8747.
- Wojcicki, M. & D. R. Brooks. 2004. Escaping the matrix: A new algorithm for phylogenetic comparative studies of co-evolution. *Cladistics* 20: 341–361.
- & ———. 2005. PACT: A simple and efficient algorithm for generating area cladograms. *J. Biogeogr.* 32: 755–774.
- Zandee, M. & M. C. Roos. 1987. Component-compatibility in historical biogeography. *Cladistics* 3: 305–332.
- Zrzavy, J. & Z. Skala. 1993. Holobionts, hybrids and cladistics. *BioSystems* 31: 127–130.